11-17-00 A/Box seq

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	B0801/7195
First Named Inventor or Application Identifier	
CASTRILLON, Diego	
Express Mail Label No.	EL711250799US
Date of Deposit	November 16, 2000

JC871 U.S. PTO 11/16/00

JC864 U.S. PTO 09/714865 11/16/00

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents

ADDRESS TO: **Box Patent Application**
Commissioner for Patents
Washington, DC 20231

1. ☐ Fee Transmittal Form
(Submit an original, and a duplicate for fee processing)

2. ✓ Specification [Total pages 51]
41 - pages description
1 - pages abstract
9 - pages claims 88 - Total claims

3. ✓ Drawing(s) *(35 USC 113)* [Total sheets 1]
☐ Informal ✓ Formal [Total drawings 1]

4. ☐ Oath or Declaration [Total pages]
a. ☐ Newly executed (original or copy)
b. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)
[Note Box 5 below]
i. ☐ DELETION OF INVENTOR(S)
Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).

5. ☐ Incorporation by Reference
(usable if Box 4b is checked)
The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

6. ☐ Microfiche Computer Program (Appendix)

7. ✓ Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
a. ✓ Computer Readable Copy
b. ✓ Paper Copy (identical to computer copy)
c. ✓ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers/cover sheet & documents(s)
9. ☐ 37 CFR 3.73(b) Statement
(when there is an assignee)
☐ Power of Attorney
10. ☐ English Translation of Document *(if applicable)*
11. ☐ Information Disclosure Statement PTO-1449
☐ Copies of IDS Citations
12. ✓ Preliminary Amendment
13. ✓ Return Receipt Postcard (MPEP 503)
(Should be specifically itemized)
14. ☐ Small Entity Statement(s)
☐ Statement filed in prior application, Status still proper and desired
15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)
16. Other: This application claims priority under Title 35 U.S.C., §119(e), of United States Provisional Application No. 60/166,394, filed November 18, 1999, entitled "COMPOSITIONS AND METHODS FOR THE IMPROVED DIAGNOSIS AND TREATMENT OF GERM CELL TUMORS," the entire contents of which are incorporated herein by reference.

17. **If a CONTINUING APPLICATION,** check appropriate box and supply the requisite information:

☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No.:

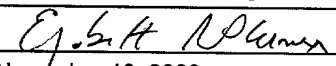
☐ Cancel in this application original claims of the prior application before calculating the filing fee.

☐ Amend the specification by inserting before the first line the sentence:

This application is a ☐ continuation ☐ divisional of application serial no. , filed , entitled , and now .

490916.1

18. CORRESPONDENCE ADDRESS					
Correspondence address below					
ATTORNEY'S NAME	Elizabeth R. Plumer, Reg. No. 36,637				
NAME	Wolf, Greenfield & Sacks, P.C.				
ADDRESS	600 Atlantic Avenue				
CITY	Boston	***STATE***	MA	***ZIP***	02210
COUNTRY	USA	***TELEPHONE***	(617) 720-3500	***FAX***	(617) 720-2441

19. SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED	
NAME	Elizabeth R. Plumer, Reg. No. 36,637
SIGNATURE	
DATE	November 16, 2000

ATTORNEY DOCKET NUMBER: B0801/7195(ERP/KA)
Express Mail Label No.: EL711250799US
Date of Deposit: November 16, 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Diego H. Castrillon
Serial No:	Not yet Assigned
Filed:	Herewith
For:	COMPOSITIONS AND METHODS FOR THE IMPROVED DIAGNOSIS OF GERM CELL TUMORS
Examiner:	Not Yet Assigned
Art Unit:	Not Yet Assigned

Box Patent Application
Commissioner for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

In the Claims:

Please cancel claims 5-11, 13-16, 18-21, 23-24, 26-34, 36-46, 48-62, 64-65, 67-79, and 81-85 prior to calculating fees, and without prejudice to future prosecution.

Please amend claims 12 and 17 as follows:

12. (Amended) An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, [2, 3, or 4] wherein the polypeptide, or fragment of the polypeptide, has germ cell specific expression.

17. (Amended) An isolated binding polypeptide which binds selectively a polypeptide encoded by the isolated nucleic acid molecule of claim 1[, 2, 3 or 4].

REMARKS

Claims 5-11, 13-16, 18-21, 23-24, 26-34, 36-46, 48-62, 64-65, 67-79, and 81-85 were cancelled without prejudice to future prosecution.

Claims 12 and 17 were amended to delete the multiple dependencies in each claim.

Respectfully Submitted,

Elizabeth R. Plumer, Reg. No. 36,637
Wolf, Greenfield & Sacks, P.C.
600 Atlantic Avenue
Boston, MA 02210-2211
(617)720-3500

B0801/7195(ERP/KA)
Date: November 16, 2000
X11/18/00

Atty Dock. No.: B0801/7195/ERP/KA
Express Mail No. EL711250799US
Date of Deposit: November 16, 2000

COMPOSITIONS AND METHODS FOR THE IMPROVED DIAGNOSIS AND TREATMENT OF GERM CELL TUMORS

Related Applications

This application claims priority under 35 USC §119(e) from U.S. Provisional Patent Application Serial No. 60/166,394, filed on November 18, 1999, entitled COMPOSITIONS AND METHODS FOR THE IMPROVED DIAGNOSIS AND TREATMENT OF GERM CELL TUMORS. The contents of the provisional application are hereby expressly incorporated by reference.

Field of the Invention

This invention relates to nucleic acids and encoded polypeptides of human *vasa*, and diagnostics and therapeutics related to medical conditions associated with such genes and polypeptides, including cancers.

Background of the Invention

Germ cells are cells that are specialized to produce haploid gametes in multicellular organisms. Germ cell tumors represent a diverse family of neoplasms affecting a wide range of patients. The great majority of testicular tumors (the most common malignancy in young men) are malignant germ cell tumors. The most common ovarian tumor, the benign or "mature" teratoma, is also of germ cell origin. Additionally, malignant ovarian germ cell tumors and germ cell tumors of various subtypes are of relatively common occurrence in children. Interestingly, benign and malignant germ cell tumors also arise in extragonadal locations (the mediastinum and central nervous system). Although the histogenesis of these extragonadal tumors is poorly understood, they exhibit similar biological behavior to their gonadal counterparts.

Malignant germ cell tumors are subdivided based on histologic appearance into pure seminomas (most common), and nonseminomas that include embryonal carcinoma, teratoma, choriocarcinoma, and yolk sac tumor. Patterns of mixed histologic appearance are also of common occurrence. Seminoma closely resembles primitive germ cells without evidence of further differentiation. Embryonal carcinoma is more anaplastic than seminoma, often with

gland-like areas, and is believed to represent early differentiation to other lineages. The remaining histologic subtypes recapitulate the various lineages of the primitive embryo, reflecting the totipotent character of germ cells. Teratoma, choriocarcinoma, and yolk sac tumors exhibit evidence of fetal, placental, and yolk sac differentiation, respectively.

Accurate tissue-based diagnosis and subtyping of germ cell tumors (i.e., seminomas vs. nonseminoma) is of paramount importance for the patient, due to differences in staging evaluation and subsequent management (See, e.g., Ch. 98 on "Testicular Cancer" in *Harrison's Principles of Internal Medicine*, 14th Edition, 1997, McGraw-Hill, Inc., N.Y. - hereinafter "Harrison's"). Yet, some germ cell tumors are difficult or even impossible to reliably diagnose or subtype due to the diverse histologic appearances of germ cell tumors and the existence of numerous other malignancies that can mimic germ cell tumors histologically. For example, clear cell carcinoma of the ovary can histologically resemble dysgerminoma (the ovarian counterparts of seminoma); mediastinal thymomas can be difficult to distinguish from germ cell tumors, and testicular lymphoma can mimic seminoma (Scully *et al*, 1998, *Armed Forces Institute of Pathology*, Washington, D.C.; Suster *et al.*, Seminars in Diagnostic Pathology, 1995, 12(1):98-104). In addition, metastases in patients with unsuspected primary germ cell tumors can be initially misdiagnosed, resulting in treatment delays.

Tumor-specific markers have been clinically useful for a variety of reasons, including accurate tissue-based diagnosis by immunohistochemistry, population-based screening, confirmation of a clinical diagnosis prior to surgery, and monitoring of patients in remission. In general, the clinical utility of a marker is directly proportional to its specificity. For example, β-hCG, a highly sensitive and specific marker for trophoblast, is indispensable in the diagnosis and clinical monitoring of patients with choriocarcinoma (it is also the basis of pregnancy tests), and α-fetoprotein (AFP) serves as a fairly sensitive marker of yolk sac differentiation. Serum assays for both AFP and β-hCG are routinely employed in the diagnosis of patients with suspected germ cell tumors to determine if non-seminomatous components are present (which would alter patient management) (Harrison's). Commercially-available AFP and β-hCG antibodies are also routinely employed in immunohistochemical assays performed by hospital laboratories on paraffin-embedded formalin-fixed tissue.

Currently available immunohistochemical markers for seminoma are relatively non-specific and no useful serum seminoma tumor markers exist, even though seminoma is the most common germ cell tumor subtype. Although placental-type alkaline phosphatase

(PLAP) is a fairly sensitive (~80%) marker of seminoma, it is far from specific, being expressed in a variety of carcinomas including the majority of ovarian carcinomas, a significant number of gastrointestinal carcinomas, and several normal tissues (Sunderland *et al.*, *Cancer Research*, 1984, 44(10):4496-4502). Largely because of this lack of specificity, alkaline phosphatase serum assays are not being utilized routinely. Furthermore, PLAP is not reliably expressed in normal germ cells (Perry *et al.*, *Human Pathology*, 1994, 25(3):235-239). The development of a more specific and sensitive marker of seminoma would represent a major advance.

The *vasa* gene was originally described in *Drosophila*, where various studies including whole-mount *in situ* staining have, reportedly, shown that *vasa* is expressed only in germ cells (of both sexes) at all stages of development, from the preblastoderm stage to primitive germ cells to gametogenesis in adults. A number of investigations have, reportedly, shown that *vasa* is not only germ-cell specific in its expression, it is absolutely required for germ cell development, and in *Drosophila*, *vasa* mutants fail to develop germ cells (Lasko *et al.*, *Nature,* 1988, 335:611-617, and SEQ ID NO:7). Subsequently, *vasa* homologues were identified in mouse, zebrafish, and *Xenopus*. In all of the foregoing species, expression was reportedly germ-cell specific and occurred throughout life. Reports on *vasa* expression in these organisms have greatly increased our knowledge of the germ-cell lineage, allowing germ-cell lineage cells to be traced back to the 4-cell morula stage (Fujiwara, *et al.*, *Proc. Nat. Acad. Scie. USA*, 1994, 6;91(25): 12258-12262; Yoon, *et al.*, *Development,* 1997,124:3157-3166; Ikenishi, *et al.*, *Dev. Grow., and Diff.,* 1997, 39:625-633).

There exists a need to identify agents that are useful in the diagnosis of tumors of germ cell origin.

These and other objects will be described in greater detail below.

Summary of the Invention

We describe herein the molecular cloning and characterization of human *vasa*, a novel molecule that has germ cell specific expression and is believed to play a determinative role in gonad development. Aberrant expression of human *vasa* has been found in patients with tumors of germ cell origin, making human *vasa* a specific marker of tumors of such origin.

The invention provides isolated nucleic acid molecules, unique fragments of those molecules, expression vectors containing the foregoing, and host cells transfected with those molecules. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies. The foregoing can be used, *inter alia,* in

the diagnosis or treatment of conditions characterized by the aberrant expression and/or the presence of mutant forms of a human *vasa* nucleic acid or polypeptide. The invention also provides methods for identifying pharmacological agents useful in the diagnosis or treatment of such conditions.

According to one aspect of the invention, isolated nucleic acid molecules that code for a human *vasa* polypeptide are provided and include: (a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleotide sequence set forth as SEQ ID NO:1 and which code for a human *vasa* polypeptide, (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and (c) complements of (a) or (b). In certain embodiments, the isolated nucleic acid molecule comprises nucleotides 1-2224 of SEQ ID NO:1. In some embodiments the isolated nucleic acid molecules are those comprising the human vasa cDNA corresponding to SEQ ID NO:15. The isolated nucleic acid molecule also can comprise a molecule which encodes the polypeptide of SEQ ID NO:2 and has human *vasa* specific expression. In certain embodiments, nucleic acids of the invention exclude nucleic acids completely composed of the nucleotide sequences of any of GenBank accession numbers listed in Table I (S75275, D14859, AB005147, Y12007, AF046043, Z81449.1, X81823, P09052, Q64060, Q61496, Q62167, O00571, P24346, P16381, O15523, AL042306, AA399611, AA398976, AA383535, AI217144, AI953070, AI025074, AI654417, AI337133, AA758412, AI969018, AA400066, AA862553, AA401568, AA316798, T85890, and T82153), or other previously published sequences as of the filing date of this application.

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) unique fragments of a nucleotide sequence set forth as SEQ ID NO:1 (of sufficient length to represent a sequence unique within the human genome), (b) complements of (a), provided that a unique fragment of (a) includes a sequence of contiguous nucleotides which is not identical to a sequence selected from the sequence group consisting of: (1) sequences having the database accession numbers of Table I, or sequences encoding a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, and SEQ ID NO:7, (2) complements of (1), and (3) fragments of (1) and (2).

In one embodiment, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous

nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

In another embodiment, the fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect, the invention provides expression vectors, and host cells transformed or transfected with such expression vectors, comprising the nucleic acid molecules described above.

According to another aspect of the invention, an isolated polypeptide is provided. The isolated polypeptide is encoded by the foregoing isolated nucleic acid molecules of the invention. In some embodiments, the isolated polypeptide is encoded by the nucleic acid of SEQ ID NO:1, giving rise to a polypeptide having the sequence of SEQ ID NO:2 that has germ cell specific expression. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing of sufficient length to represent a sequence unique within the human genome, and identifying with a polypeptide that has germ cell specific expression, provided that the fragment excludes: (i) a sequence of contiguous amino acids selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, and SEQ ID NO:7, and (ii) a sequence of contiguous amino acids encoded by an isolated nucleic acid having a nucleotide sequence with a GenBank database accession number selected from the group consisting of S75275, D14859, AB005147, Y12007, AF046043, Z81449.1, X81823, P09052, Q64060, Q61496, Q62167, O00571, P24346, P16381, O15523, AL042306, AA399611, AA398976, AA383535, AI217144, AI953070, AI025074, AI654417, AI337133, AA758412, AI969018, AA400066, AA862553, AA401568, AA316798, T85890, and T82153. (i.e., as described in Table I). In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated binding polypeptides which selectively bind a polypeptide encoded by the foregoing isolated nucleic acid molecules of the invention are provided. Preferably the isolated binding polypeptides selectively bind a polypeptide which comprises the sequence of SEQ ID NO:2, SEQ ID NO:9, SEQ ID NO:10, or fragments thereof. In preferred embodiments, the isolated binding polypeptides include

antibodies and fragments of antibodies (e.g., Fab, $F(ab')_2$, Fd and antibody fragments which include a CDR3 region which binds selectively to the human *vasa* polypeptide). In certain embodiments, the antibodies are human.

The invention also contemplates kits comprising a package including assays for human *vasa* epitopes, human *vasa* nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention, for comparing the level of human *vasa* polypeptides or human *vasa* nucleic acids in a test sample to the level in a control sample. This comparison can be used to assess in a subject the presence of a tumor of germ cell origin. The kits may also include assays for other known genes, and expression products thereof, associated with cancers (e.g., β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP -glial fibrillary acidic protein, synaptophysin, chromogranin, cytokeratin isoforms such as 7 and 20, anti-keratin markers such as AE1/AE3 and CAM5.2, etc.).

Another aspect of the invention is a method for determining the level of *vasa* expression in a subject. Expression is defined either as *vasa* mRNA expression or *vasa* polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and northern blotting for measuring mRNA expression, and monoclonal *vasa* antibodies or polyclonal *vasa* antisera as reagents to measure (or characterize) *vasa* polypeptide expression. In certain embodiments, test samples such as tissue (e.g., biopsy) samples, and biological fluids such as blood, are used as test samples. In some embodiments, fine-needle aspirates can also be used as sources of a test sample. *Vasa* expression in a test sample of a subject is compared to *vasa* expression in control sample to, e.g., assess the presence or absence, or stage of a tumor of germ cell origin in a subject.

According to another aspect, a method of detecting a tumor of germ cell origin, is provided. The method involves detecting *vasa* expression in an extragonadal test sample obtained from a subject. *Vasa* expression in the extragonadal test sample is indicative of the presence of a tumor of germ cell origin in the subject. *Vasa* expression and methods of measuring *vasa* expression are as described in any of the foregoing embodiments. In some embodiments, the subject has not previously been diagnosed as having a tumor of germ cell origin or a predisposition thereto [e.g., to detect a metastasis in a subject of previously

undiagnosed cancer]. In certain embodiments, the subject has a clinical diagnosis of a tumor of germ cell origin and the method is to confirm the clinical diagnosis, monitor remission of the tumor, or stage the tumor.

According to yet another aspect, a method of detecting a tumor of germ cell origin, is provided. The method involves detecting *vasa* overexpression in a test sample obtained from a subject. *Vasa* overexpression in the test sample as compared to a control is indicative of a tumor of germ cell origin in the subject. *Vasa* expression and methods of measuring *vasa* expression are as described in any of the foregoing embodiments. In certain embodiments, the tumor can be a testicular tumor (e.g., a seminoma), an ovarian tumor (e.g., a dysgerminoma or a teratoma), or a tumor of an extragonadal tissue(e.g., a mediastinal tumor or an intracranial tumor). In some embodiments, the method can further comprise detecting expression of a tumor-specific agent other than a *vasa* molecule (nucleic acid or polypeptide) in the test sample. Tumor-specific agents other than a *vasa* molecule include, but are not limited to, β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP -glial fibrillary acidic protein, synaptophysin, chromogranin, cytokeratin isoforms such as 7 and 20, and anti-keratin markers such as AE1/AE3 and CAM5.2. Preferred subjects are as described in any of the foregoing embodiments.

According to another aspect, a method of subtyping tumors of germ cell origin is provided. The method involves detecting *vasa* expression in a test sample of a known or suspected germ cell origin tumor obtained from a subject. *Vasa* overexpression in the test sample as compared to a control is indicative of a seminoma in the subject, whereas absence of *vasa* expression in the test sample as compared to a control is indicative of a nonseminoma in the subject. *Vasa* expression, methods of measuring *vasa* expression, tumor types, and subjects, are as described in any of the foregoing embodiments. Nonseminomas include embryonal carcinoma, teratoma, choriocarcinoma, yolk sac tumor, or combinations of the foregoing. In certain embodiments, the subject has a clinical diagnosis of a tumor of mixed histologic appearance. In some embodiments, and preferably (but not exclusively) in the absence of *vasa* expression in the test sample, the method can further comprise detecting expression of a tumor-specific agent other than a *vasa* molecule in the test sample. Tumor-specific agents other than a *vasa* molecule are as described in any of the foregoing embodiments.

According to another aspect, a method of distinguishing a tumor of germ cell origin from a non-germ cell tumor, is provided. The method involves detecting expression of a *vasa*

molecule in a test sample, wherein expression of the *vasa* molecule is indicative of a tumor of a germ cell origin and absence of expression of the *vasa* molecule is indicative of a non-germ cell tumor. In certain embodiments, the non-germ cell tumor resembles histologically a tumor of germ cell origin. Non-germ cell tumors include, but are not limited to, clear cell carcinoma of the ovary (can resemble dysgerminoma), a mediastinal thymoma, and a testicular lymphoma.

According to a further aspect, a method for treating a tumor of germ cell origin in a subject, is provided. The method involves administering to a subject in need of such treatment an agent that inhibits *vasa* expression in a germ cell of the subject in an effective amount to inhibit *vasa* expression and inhibit the growth and/or proliferation of the tumor of germ cell origin in the subject. In a preferred embodiment, the agent is a *vasa* antisense nucleic acid. In certain embodiments, the method further comprises co-administering an anti-cancer agent.

According to another aspect of the invention, methods for preparing medicaments useful in the treatment of a tumor of germ cell origin, are also provided.

The present invention thus involves, in several aspects, human *vasa* polypeptides, isolated nucleic acids encoding those polypeptides, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as therapeutics and diagnostics relating thereto.

These and other objects of the invention will be described in further detail in connection with the detailed description of the invention.

Brief Description of the Drawings

Figure 1 depicts a kit comprising an agent of the invention (e.g., anti-human *vasa* Abs, human *vasa* epitopes, etc.), a control agent, and instructions for utilizing such agents in diagnostic or therapeutic applications.

Brief Description of the Sequences

SEQ ID NO:1 is the nucleotide sequence of the human *vasa* cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human *vasa* cDNA (SEQ ID NO:1).

SEQ ID NO:3 is the amino acid sequence of the *Mus Musculus vasa* cDNA.

SEQ ID NO:4 is the amino acid sequence of the *Rattus Norvegicus vasa* cDNA.

SEQ ID NO:5 is the amino acid sequence of the *Xenopus Laevis vasa* cDNA.

SEQ ID NO:6 is the amino acid sequence of the *Danio Reio vasa* cDNA.

SEQ ID NO:7 is the amino acid sequence of the *Drosophila Melanogaster vasa* cDNA.

SEQ ID NO:8 is the nucleotide sequence of a synthetic oligonucleotide primer used in the cloning of the human *vasa* cDNA.

SEQ ID NO:9 is the amino acid sequence of a human-specific *vasa* epitope used in the generation of anti-human *vasa* antibodies.

SEQ ID NO:10 is the amino acid sequence of a human-specific *vasa* epitope used in the generation of anti-human *vasa* antibodies.

SEQ ID NO:11 is the nucleotide sequence of a synthetic human *vasa* oligonucleotide.

SEQ ID NO:12 is the nucleotide sequence of a synthetic human *vasa* oligonucleotide.

SEQ ID NO:13 is the nucleotide sequence of a synthetic oligonucleotide primer used in the detection of human *vasa* expression using PCR in combination with the synthetic oligonucleotide primer of SEQ ID NO:14.

SEQ ID NO:14 is the nucleotide sequence of a synthetic oligonucleotide primer used in the detection of human *vasa* expression using PCR in combination with the synthetic oligonucleotide primer of SEQ ID NO:13.

SEQ ID NO:15 is the nucleotide sequence of the largest open reading frame of the human *vasa* cDNA of SEQ ID NO:1, encoding for the polypeptide of SEQ ID NO:2.

Detailed Description of the Invention

One aspect of the invention involves the cloning of a cDNA encoding human *vasa*. Human *vasa* according to the invention is an isolated nucleic acid molecule that comprises a nucleic acid molecule of SEQ ID NO:1, and codes for a protein that is specifically expressed in the gonads and is believed to play an essential role in gonad development. According to the invention, aberrant expression of human *vasa* has been found in patients with tumors of germ cell origin, making human *vasa* a specific marker for tumors of such origin. The sequence of the human *vasa* cDNA is presented as SEQ ID NO:1, and the predicted amino acid sequence of this cDNA's encoded protein product is presented as SEQ ID NO:2.

As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments human *vasa* and human subjects are preferred.

The invention thus involves in one aspect an isolated human *vasa* polypeptide, the cDNA encoding this polypeptide, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as diagnostics and therapeutics relating thereto.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides, the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

According to the invention, isolated nucleic acid molecules that code for a human *vasa* polypeptide include: (a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleic acid of SEQ ID NO:1 and which code for a human *vasa* polypeptide, (b) deletions, additions and substitutions of (a) which code for a respective human *vasa* polypeptide, (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and (d) complements of (a), (b) or (c). In certain embodiments, nucleic acids of the invention exclude nucleic acids completely composed of the nucleotide sequences of any of GenBank accession numbers listed in Table I (S75275, D14859, AB005147, Y12007, AF046043, Z81449.1, X81823, P09052, Q64060, Q61496, Q62167, O00571, P24346, P16381, O15523, AL042306, AA399611, AA398976, AA383535, AI217144, AI953070, AI025074, AI654417, AI337133, AA758412, AI969018, AA400066, AA862553, AA401568, AA316798, T85890, and T82153), or other previously published sequences as of the filing date of this application.

Homologs and alleles of the human *vasa* nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for human *vasa* polypeptides and which hybridize to a nucleic acid molecule consisting of the coding region of SEQ ID NO:1, under stringent conditions. The

term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH_2PO_4(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of human *vasa* nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 75% nucleotide identity and/or at least 90% amino acid identity to SEQ ID NO:1 and SEQ ID NO:2, respectively. In some instances homologs and alleles will share at least 90% nucleotide identity and/or at least 95% amino acid identity and in still other instances will share at least 95% nucleotide identity and/or at least 99% amino acid identity. Complements of the foregoing nucleic acids also are embraced by the invention. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland). Exemplary tools include the heuristic algorithm of Altschul SF, et al., (*J Mol Biol*, 1990, 215:403-410), also known as BLAST. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using public (EMBL, Heidelberg, Germany) and commercial (e.g., the MacVector sequence analysis software from Oxford Molecular Group/Genetics Computer Group, Madison, WI). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for human *vasa* related genes, such as homologs and alleles of human *vasa*, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphoimager plate to detect the radioactive signal.

Given that the expression of the human *vasa* gene is abundant in certain human tissues, and given the teachings herein of a human *vasa* cDNA clone encoding for the complete *vasa* polypeptide, allelic human *vasa* sequences can be isolated from cDNA libraries prepared from one or more of the tissues in which human *vasa* expression is abundant (i.e., in the gonads), using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating human *vasa* polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of SEQ ID NO:1 or SEQ ID NO:15, or complements thereof. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the human *vasa* nucleic acids defined above. Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome. Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers listed in Table I (S75275, D14859, AB005147, Y12007, AF046043, Z81449.1, X81823, P09052, Q64060, Q61496, Q62167, O00571, P24346, P16381, O15523, AL042306, AA399611, AA398976, AA383535, AI217144, AI953070,

AI025074, AI654417, AI337133, AA758412, AI969018, AA400066, AA862553, AA401568, AA316798, T85890, and T82153), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the human *vasa* polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of human *vasa* nucleic acids and polypeptides respectively.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1 or SEQ ID NO:15 and complements will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 2224, or SEQ ID NO:15 beginning at nucleotide 1 and ending at nucleotide 2172, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique fragment to

selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a human *vasa* polypeptide, to decrease human *vasa* activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1 or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nat. Med.* 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the

antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1. Similarly, antisense to allelic or homologous human *vasa* cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical

preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding human *vasa* polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for human *vasa* proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as E.coli and mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time

per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences.

The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding human *vasa* polypeptide or fragment or variant thereof. That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV and pcDNA3.1 (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1α, which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier *et al.*, in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, human *vasa* cDNA sequence containing expression vectors, to transfect host cells and cell lines, be these prokaryotic (e.g., E. coli), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide

variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of human *vasa* gene "knock-outs" in cells and in animals, providing materials for studying certain aspects of human *vasa* activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing human *vasa* nucleic acids, and include the polypeptide of SEQ ID NO:2 and unique fragments thereof. Such polypeptides are useful, for example to generate antibodies, as components of an immunoassay, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of an human *vasa* polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2 will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, 474 amino acids long). Virtually any segment of SEQ ID NO:2, excluding the ones that share identity with it (the polypeptides having amino acid sequences as set forth in SEQ ID NOs: 3, 4, 5, 6, and 7) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include interaction with antibodies, interaction with other polypeptides or fragments thereof, selective binding to receptors, tissue specific expression, etc. One important activity is the ability to act as a signature for identifying the polypeptide. Another is the ability to complex with HLA and to provoke in a human an immune response. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to

selectively distinguish the sequence of interest from non-family members or from family members of other species. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the human *vasa* polypeptides described above. As used herein, a "variant" of a human *vasa* polypeptide is a polypeptide which contains one or more modification to the primary amino acid sequence of a human *vasa* polypeptide. Modifications which create a human *vasa* polypeptide variant are typically made to the nucleic acid which encodes the human *vasa* polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate an activity of a human *vasa* polypeptide; 2) enhance a property of a human *vasa* polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a human *vasa* polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a human *vasa* polypeptide receptor or other molecule (e.g., heparin). Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the human *vasa* amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant human *vasa* polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary a only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

Variants can include human *vasa* polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a human *vasa* polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

Mutations of a nucleic acid which encodes a human *vasa* polypeptide preferably preserve the amino acid reading frame of the coding sequence, and preferably do not create

regions in the nucleic acid which are likely to hybridize to form secondary structures, such a hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant human *vasa* polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., E. coli, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a human *vasa* gene or cDNA clone to enhance expression of the polypeptide.

The skilled artisan will realize that conservative amino acid substitutions may be made in human *vasa* polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the human *vasa* polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the human *vasa* polypeptides include conservative amino acid substitutions of SEQ ID NO:2. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

Thus functionally equivalent variants of human *vasa* polypeptides, i.e., variants of human *vasa* polypeptides which retain the function of the natural human *vasa* polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of human *vasa* polypeptides to produce functionally equivalent variants of human *vasa* polypeptides typically are made by alteration of a nucleic acid encoding human *vasa* polypeptides (SEQ ID NOs:1, 15). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel

(Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a human *vasa* polypeptide. The activity of functionally equivalent fragments of human *vasa* polypeptides can be tested by cloning the gene encoding the altered human *vasa* polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered human *vasa* polypeptide, and testing for a functional capability of the human *vasa* polypeptides as disclosed herein (e.g., germ cell specific expression, etc.).

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of human *vasa* polypeptides. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated human *vasa* molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of human *vasa* mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce human *vasa* polypeptides. Those skilled in the art also can readily follow known methods for isolating human *vasa* polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from human *vasa* polypeptides. A dominant negative polypeptide is an inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative catalytically-inactive kinase which interacts normally with target proteins but does not phosphorylate the target proteins can reduce phosphorylation of the target proteins in response to a cellular signal. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

The isolation of the human *vasa* cDNA also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of human *vasa*. These methods involve determining expression of the human *vasa* gene, and/or human *vasa* polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the secreted human *vasa* protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to human *vasa* polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 3, 4, 5, 6, and 7, and/or the nucleic acids having nucleotide sequences with GenBank accession numbers as those described in Table I, and/or polypeptides having amino acid sequences with GenBank accession numbers as those described in Table I..

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) *The Experimental Foundations of Modern Immunology,* Wiley & Sons, Inc., New York; Roitt, I. (1991) *Essential Immunology*, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated

an $F(ab')_2$ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for $F(ab')_2$, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric $F(ab')_2$ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by

homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves polypeptides of numerous size and type that bind specifically to human *vasa* polypeptides, and complexes of both human *vasa* polypeptides and their binding partners. These polypeptides may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the human *vasa* polypeptide or a complex of human *vasa* and a binding partner. This process can be repeated through several cycles of reselection of phage that bind to the human *vasa* polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the human *vasa* polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the human *vasa* polypeptides. Thus, the human *vasa* polypeptides of the invention, or a fragment thereof, or complexes of human *vasa* and a binding partner can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding partners of the human *vasa* polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of human *vasa* and for other purposes that will be apparent to those of ordinary skill in the art.

A human *vasa* polypeptide, or a fragment thereof, also can be used to isolate their native binding partners. Isolation of binding partners may be performed according to well-known methods. For example, isolated human *vasa* polypeptides (that include human

vasa phosporylated polypeptides) can be attached to a substrate, and then a solution suspected of containing an human *vasa* binding partner may be applied to the substrate. If the binding partner for human *vasa* polypeptides is present in the solution, then it will bind to the substrate-bound human *vasa* polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for human *vasa*, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-human *vasa* antibodies. In the case of nucleic acid detection, pairs of primers for amplifying human *vasa* nucleic acids can be included. The preferred kits would include controls such as known amounts of nucleic acid probes, human *vasa* epitopes (such as human *vasa* expression products) or anti-human *vasa* antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize the presence of a tumor based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with human *vasa* protein and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention, generally designated by the numeral 11, is illustrated in Figure 1. Kit 11 is comprised of the following major elements: packaging 15, an agent of the invention 17 (e.g., an anti-human *vasa* antibody), a control agent 19 (e.g., a human *vasa* epitope), and instructions 21. Packaging 15 is a box-like structure for holding a vial (or number of vials) containing an agent of the invention 17, a vial (or number of vials) containing a control agent 19, and instructions 21. Individuals skilled in the art can readily modify packaging 15 to suit individual needs.

The invention also provides methods to measure the level of human *vasa* expression in a subject. This can be performed by first obtaining a test sample from the subject. The test sample can be tissue, biological fluid, or both (as, for example, in a fine needle aspirate). Tissues include brain, heart, serum, breast, colon, bladder, uterus, prostate, stomach, testis, ovary, pancreas, pituitary gland, adrenal gland, thyroid gland, salivary gland, mammary gland, kidney, liver, intestine, spleen, thymus, bone marrow, trachea, and lung. In certain embodiments, test samples originate from brain, testis, breast and prostate tissues, and biological fluids include blood, saliva, semen, fellopian fluid, and urine. Both invasive and

non-invasive techniques can be used to obtain such samples and are well documented in the art. "Vasa expression," as used herein, is used interchangeably with "*vasa* molecule expression," and refers to a *vasa* nucleic acid or a *vasa* peptide (including polypeptide and protein) expression. At the molecular level both PCR and Northern blotting can be used to determine the level of human *vasa* mRNA using products of this invention described earlier, and protocols well known in the art that are found in references which compile such methods. At the protein level, human *vasa* expression can be determined using either polyclonal anti-human *vasa* sera or monoclonal antibodies in combination with standard immunological assays. The preferred methods of the invention compare the measured level of human *vasa* expression of the test sample to a control having a known human *vasa* expression (e.g., to assess the presence or absence, or stage of a tumor of germ cell origin in a subject). A control human *vasa* expression can include a known amount of a nucleic acid probe, a human *vasa* epitope (such as a human *vasa* expression product), or a sample from the same tissue (fluid or aspirate) of a subject with a 'normal' (control) level of human *vasa* expression.

The invention also embraces methods of detecting tumors of germ cell origin. According to this aspect of the invention, the test sample can be of extragonadal (nongonadal) or gonadal tissue origin. When the test sample is of extragonadal tissue origin, the method involves detecting *vasa* expression in a test sample obtained from an extragonadal tissue of a subject. *Vasa* expression in the extragonadal test sample is indicative of the presence of a tumor of germ cell origin in the subject. The preferred methods of the invention compare the measured level of human *vasa* expression in the extragonadal test sample to the level of *vasa* expression in a control sample, preferably from the same extragonadal tissue of a 'normal' (control) subject. In general, extragonadal test samples from 'normal' subjects contain undetectable expression levels of a *vasa* molecules; test samples having detectable *vasa* expression levels are indicative of the presence of a tumor of germ cell origin in the subject.

When the test sample is obtained from a gonadal tissue of the subject (e.g., ovaries, testis, etc.), the method involves detecting *vasa* overexpression in the test sample, since *vasa* molecules are typically expressed at a baseline level corresponding to a 'normal' level in control gonadal tissues. Examination of gonadal tissue test samples typically involves examination of histologic samples (e.g., biopsy tissue slides). Histologic samples are commonly stained with agents that reveal tissue/cell morphology. Preferably, the skilled artisan selects histologic samples (e.g., from a gonadal tissue biopsy) that include tissue areas depicting both 'abnormal' and 'normal' morphologic appearance. The tissue areas with 'normal' appearance serve as an internal negative control (a preferred control). Alternatively,

or additionally, gonadal tissues from 'normal' subjects can be used as a negative control. To detect the presence of a tumor of germ cell origin in gonadal tissue, the test samples are compared to the control samples (e.g., negative controls as discussed above or positive controls, i.e., samples of gonadal tissue having tumors of known germ cell origin) and *vasa* overexpression is determined. "Overexpression," as used in reference to a histologic test sample, refers to a statistically significant increase in vasa expression per unit surface area compared to a negative control sample. Therefore, *vasa* overexpression in the gonadal test sample as compared to a negative control sample, is indicative of the presence of a tumor of germ cell origin in the subject. In some embodiments, the subject has not previously been diagnosed as having a tumor of germ cell origin or a predisposition thereto (e.g., to detect a metastasis in a subject of previously undiagnosed cancer). In certain embodiments, the subject has a clinical diagnosis of a tumor of germ cell origin and the method is to confirm the clinical diagnosis, monitor remission of the tumor, or stage the tumor. In some embodiments, the tumor can be a testicular tumor (e.g., a seminoma), an ovarian tumor (e.g., a dysgerminoma or a teratoma), or a tumor of an extragonadal tissue (e.g., a mediastinal tumor or an intracranial tumor). In preferred embodiments, when the test sample examined is of gonadal origin, the method can further comprise detecting expression of a tumor-specific agent other than a *vasa* molecule (nucleic acid or polypeptide) in the test sample. Tumor-specific agents other than a *vasa* molecule include, but are not limited to, β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP -glial fibrillary acidic protein, synaptophysin, chromogranin, cytokeratin isoforms such as 7 and 20, and anti-keratin markers such as AE1/AE3 and CAM5.2.

According to another aspect, a method of subtyping tumors of germ cell origin is provided. The method involves detecting *vasa* expression in a test sample of a known or suspected germ cell origin tumor obtained from a subject. *Vasa* overexpression in the test sample as compared to a control is indicative of a seminoma in the subject, whereas absence of *vasa* expression in the test sample as compared to a control can be indicative of a nonseminoma in the subject. Nonseminomas include embryonal carcinoma, teratoma, choriocarcinoma, yolk sac tumor, or combinations of the foregoing. In certain embodiments, the subject has a tumor of mixed histologic appearance. In some embodiments, and preferably (but not exclusively) in the absence of *vasa* expression in the test sample, the method can further comprise detecting expression of a tumor-specific agent other than a *vasa* molecule in

the test sample. Tumor-specific agents other than a *vasa* molecule are as described elsewhere herein.

According to another aspect, a method of distinguishing a tumor of germ cell origin from a non-germ cell tumor, is provided. The method involves detecting expression of a *vasa* molecule in a test sample, wherein expression of the *vasa* molecule is indicative of a tumor of a germ cell origin, and absence of expression of the *vasa* molecule is indicative of a non-germ cell tumor. In certain embodiments, the non-germ cell tumor resembles histologically a tumor of germ cell origin. Non-germ cell tumors include, but are not limited to, clear cell carcinoma of the ovary (can resemble dysgerminoma), a mediastinal thymoma, and a testicular lymphoma.

In preferred embodiments of the invention tumors aberrantly expressing human *vasa* can include: biliary tract cancer; brain cancer, including glioblastomas and medulloblastomas; breast cancer; cervical cancer; choriocarcinoma; colon cancer; endometrial cancer; esophageal cancer; gastric cancer; hematological neoplasms, including acute lymphocytic and myelogenous leukemia; multiple myeloma; AIDS associated leukemias and adult T-cell leukemia lymphoma; intraepithelial neoplasms, including Bowen's disease and Paget's disease; liver cancer; lung cancer; lymphomas, including Hodgkin's disease and lymphocytic lymphomas; neuroblastomas; oral cancer, including squamous cell carcinoma; ovarian cancer, including those arising from epithelial cells, stromal cells, germ cells and mesenchymal cells; pancreas cancer; prostate cancer; colorectal cancer; sarcomas, including leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma and osteosarcoma; skin cancer, including melanoma, Kaposi's sarcoma, basocellular cancer and squamous cell cancer; testicular cancer, including germinal tumors (seminoma, non-seminoma[teratomas, choriocarcinomas]), stromal tumors and germ cell tumors; thyroid cancer, including thyroid adenocarcinoma and medullar carcinoma; and renal cancer including adenocarcinoma and Wilms tumor. Aberrant expression of a *vasa* molecule in all of the foregoing tumors is indicative that such tumors originated from cells of germ cell origin, which cells either through development or metastases became part of the tissue presently characterized as tumorous. In preferred embodiments, aberrant expression is overexpression of a human *vasa* molecule.

Human *vasa* polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced human *vasa* polypeptides include chimeric proteins comprising a fusion of a human *vasa* protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the

human *vasa* polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a human *vasa* polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of human *vasa* polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β-galactosidase activity, luciferase activity, and the like. For cell free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc). or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a human *vasa* binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

The invention provides human *vasa*-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, human *vasa*-specific pharmacological agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered human *vasa* binding characteristics. Novel human *vasa*-specific binding agents include human *vasa*-specific antibodies, cell surface receptors, and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

In general, the specificity of human *vasa* binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a human *vasa* polypeptide preferably have binding equilibrium constants of at least about 10^7 M^{-1}, more preferably at least about 10^8 M^{-1}, and most preferably at least about 10^9 M^{-1}. The wide

variety of cell based and cell free assays may be used to demonstrate human *vasa*-specific binding. Cell based assays include one, two and three hybrid screens, assays in which human *vasa*-mediated transcription is inhibited or increased, etc. Cell free assays include human *vasa*-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind human *vasa* polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

According to a further aspect, a method for treating a tumor of germ cell origin in a subject, is provided. The method involves administering to a subject in need of such treatment an agent that inhibits *vasa* expression in a germ cell of the subject in an effective amount to inhibit *vasa* expression and inhibit the growth and/or proliferation of the tumor of germ cell origin in the subject. In a preferred embodiment, the agent is a *vasa* antisense nucleic acid. In certain embodiments, the method further comrises co-administering an anti-cancer agent.

Anti-cancer agents include, but are not limited to: Acivicin; Aclarubicin; Acodazole Hydrochloride; Acronine; Adozelesin; Aldesleukin; Altretamine; Ambomycin; Ametantrone Acetate; Aminoglutethimide; Amsacrine; Anastrozole; Anthramycin; Asparaginase; Asperlin; Azacitidine; Azetepa; Azotomycin; Batimastat; Benzodepa; Bicalutamide; Bisantrene Hydrochloride; Bisnafide Dimesylate; Bizelesin; Bleomycin Sulfate; Brequinar Sodium; Bropirimine; Busulfan; Cactinomycin; Calusterone; Caracemide; Carbetimer; Carboplatin; Carmustine; Carubicin Hydrochloride; Carzelesin; Cedefingol; Chlorambucil; Cirolemycin; Cisplatin; Cladribine; Crisnatol Mesylate; Cyclophosphamide; Cytarabine; Dacarbazine; Dactinomycin; Daunorubicin Hydrochloride; Decitabine; Dexormaplatin; Dezaguanine; Dezaguanine Mesylate; Diaziquone; Docetaxel; Doxorubicin; Doxorubicin Hydrochloride; Droloxifene; Droloxifene Citrate; Dromostanolone Propionate; Duazomycin; Edatrexate; Eflornithine Hydrochloride; Elsamitrucin; Enloplatin; Enpromate; Epipropidine; Epirubicin Hydrochloride; Erbulozole; Esorubicin Hydrochloride; Estramustine; Estramustine Phosphate Sodium; Etanidazole; Etoposide; Etoposide Phosphate; Etoprine; Fadrozole Hydrochloride; Fazarabine; Fenretinide; Floxuridine; Fludarabine Phosphate; Fluorouracil; Flurocitabine; Fosquidone; Fostriecin Sodium; Gemcitabine; Gemcitabine Hydrochloride; Hydroxyurea; Idarubicin Hydrochloride; Ifosfamide; Ilmofosine; Interferon Alfa-2a; Interferon Alfa-2b; Interferon Alfa-n1; Interferon Alfa-n3; Interferon Beta-I a; Interferon Gamma-I b; Iproplatin; Irinotecan Hydrochloride; Lanreotide Acetate; Letrozole; Leuprolide Acetate; Liarozole Hydrochloride; Lometrexol Sodium; Lomustine; Losoxantrone Hydrochloride; Masoprocol; Maytansine; Mechlorethamine Hydrochloride; Megestrol Acetate; Melengestrol Acetate;

Melphalan; Menogaril; Mercaptopurine; Methotrexate; Methotrexate Sodium; Metoprine; Meturedepa; Mitindomide; Mitocarcin; Mitocromin; Mitogillin; Mitomalcin; Mitomycin; Mitosper; Mitotane; Mitoxantrone Hydrochloride; Mycophenolic Acid; Nocodazole; Nogalamycin; Ormaplatin; Oxisuran; Paclitaxel; Pegaspargase; Peliomycin; Pentamustine; Peplomycin Sulfate; Perfosfamide; Pipobroman; Piposulfan; Piroxantrone Hydrochloride; Plicamycin; Plomestane; Podofilox; Porfimer Sodium; Porfiromycin; Prednimustine; Procarbazine Hydrochloride; Puromycin; Puromycin Hydrochloride; Pyrazofurin; Riboprine; Rogletimide; Safingol; Safingol Hydrochloride; Semustine; Simtrazene; Sparfosate Sodium; Sparsomycin; Spirogermanium Hydrochloride; Spiromustine; Spiroplatin; Streptonigrin; Streptozocin; Sulofenur; Talisomycin; Taxotere; Tecogalan Sodium; Tegafur; Teloxantrone Hydrochloride; Temoporfin; Teniposide; Teroxirone; Testolactone; Thiamiprine; Thioguanine; Thiotepa; Tiazofurin; Tirapazamine; Topotecan Hydrochloride; Toremifene Citrate; Trestolone Acetate; Triciribine Phosphate; Trimetrexate; Trimetrexate Glucuronate; Triptorelin; Tubulozole Hydrochloride; Uracil Mustard; Uredepa; Vapreotide; Verteporfin; Vinblastine Sulfate; Vincristine Sulfate; Vindesine; Vindesine Sulfate; Vinepidine Sulfate; Vinglycinate Sulfate; Vinleurosine Sulfate; Vinorelbine Tartrate; Vinrosidine Sulfate; Vinzolidine Sulfate; Vorozole; Zeniplatin; Zinostatin; Zorubicin Hydrochloride.

The *vasa* inhibitory agents of the invention (e.g., *vasa* antisense molecules, *vasa* binding molecules) are administered in effective amounts. An effective amount is a dosage of the *vasa* inhibitory agent sufficient to provide a medically desirable result. The effective amount will vary with the particular condition being treated, the age and physical condition of the subject being treated, the severity of the condition, the duration of the treatment, the nature of the concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. For example, in connection with treating a tumor of germ cell origin in a subject, an effective amount is that amount which inhibits or reduces growth and/or proliferation of the tumor in the subject. Thus, it will be understood that the *vasa* inhibitory agents of the invention can be used to treat the above-noted conditions according to the preferred modes of administration described below. It is preferred generally that a maximum dose be used, that is, the highest safe dose according to sound medical judgment.

A subject, as used herein, refers to a human with a tumor of germ cell origin.

A *vasa* inhibitory agent of the invention may be administered alone or as part of a pharmaceutical composition. Such a pharmaceutical composition may include the *vasa* inhibitory agent in combination with any standard physiologically and/or pharmaceutically

acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the *vasa* inhibitory agent in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically-acceptable carrier" as used herein means one or more compatible solid or liquid filler, diluents or encapsulating substances which are suitable for administration into a human or other animal. The term "carrier" denotes an organic or inorganic ingredient, natural or synthetic, with which the active ingredient is combined to facilitate the application. The components of the pharmaceutical compositions also are capable of being co-mingled with the molecules of the present invention, and with each other, in a manner such that there is no interaction which would substantially impair the desired pharmaceutical efficacy. Pharmaceutically acceptable further means a non-toxic material, that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils.

Compositions suitable for parenteral administration conveniently comprise sterile aqueous and non-aqueous preparations of the *vasa* inhibitory agents of the invention. This aqueous preparation may be formulated according to known methods using suitable dispersing or wetting agents and suspending agents. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate, and including synthetic mono- or di-glycerides. The sterile injectable preparation also may be a sterile injectable solution or suspension in a non-toxic parenterally-acceptable diluent or solvent, for example, as a solution in 1,3-butane diol. In addition, fatty acids such as oleic acid may be used in the preparation of injectables. Carrier formulations suitable for oral, subcutaneous, intravenous, intramuscular, etc. administrations can be found in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, PA. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

Parenteral administration routes are preferred according to the present invention. Any local parenteral administration that is medically acceptable, meaning any local administration that produces effective levels of the active compounds without causing clinically

unacceptable adverse effects can be used to deliver agents of the invention. Delivery by injection into the affected gonadal tissue is preferred.

The pharmaceutical preparations, as described above, are administered in effective amounts. The effective amount will depend upon the mode of administration, the particular condition being treated and the desired outcome. It will also depend upon, as discussed above, the stage of the condition, the age and physical condition of the subject, the nature of concurrent therapy, if any, and like factors well known to the medical practitioner. For therapeutic applications, it is that amount sufficient to achieve a medically desirable result. In some cases this is a decrease in germ cell growth and/or proliferation.

Generally, doses of active compounds of the present invention would be from about 0.01 mg/kg per day to 1000 mg/kg per day. It is expected that doses ranging from 50-500 mg/kg will be suitable.

The pharmaceutical compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well-known in the art of pharmacy. All methods include the step of bringing the *vasa* inhibitory agents of the invention into association with a carrier which constitutes one or more accessory ingredients. In general, the compositions are prepared by uniformly and intimately bringing the *vasa* inhibitory agents into association with a liquid carrier, a finely divided solid carrier, or both, and then, if necessary, shaping the product.

Compositions suitable for parenteral administration (e.g., by direct/local injection) include suspensions in aqueous liquids or non-aqueous liquids such as a syrup, elixir or an emulsion.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the *vasa* inhibitory agents of the invention, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include the above-described polymeric systems, as well as polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones, polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sylastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples

include, but are not limited to: (a) erosional systems in which the *vasa* inhibitory agent is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, are used herein, means that the implant is constructed and arranged to delivery therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The *vasa* inhibitory agents of the invention may be administered alone or in combination (co-administered) with the above-described drug therapies by any conventional route, including injection, repeated injection, topical application, etc., over time. The administration may, for example, be oral, intraperitoneal, intramuscular, intra-cavity, subcutaneous, intravenous or transdermal for the co-administered anti-cancer agent . When using the *vasa* inhibitory agents of the invention, direct administration to the affected site (e.g., ovaries, testis, etc.) such as administration by injection, is preferred.

The term "co-administered," means administered substantially simultaneously with another anti-cancer agent. By substantially simultaneously, it is meant that a *vasa* inhibitory agent of the invention is administered to the subject close enough in time with the administration of the anti-cancer agent. The anti-cancer agent may be present in a different formulation than the *vasa* inhibitory agent of the invention, or it may be part of the same formulation (and therefore be administered locally together with the agent of the invention).

The co-administered agent can act cooperatively, additively or synergistically with a *vasa* inhibitory agent of the invention to produce a desired effect, for example, inhibition of the tumor of germ cell origin. Since germ cells are relatively 'dispensible' (i.e., necessary for reproduction but not necessary for survival), as long as the agent is delivered to the gonads, normal (nontumorous) germ cells can also be targeted with the therapeutic agents of the invention. The anti-cancer agent is administered in effective amounts. Such amounts maybe less than these sufficient to provide a therapeutic benefit when the agent is administered alone and not in combination with a *vasa* inhibitory agent of the invention. A person of ordinary skill in the art would be able to determine the effective amounts needed.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

Example 1: *Cloning of human vasa gene*

Expressed sequence tags (EST's) corresponding to the human *vasa* gene were identified in a publically available database maintained by the National Center for Biotechnology Information (NCBI) using the BLAST program. The DNA sequence corresponding to the mouse *vasa* cDNA (Genbank accession #2500525) was the query sequence; default search parameters were employed. Several EST's (zt93a07.sl- Genbank# AA399611; zt93a07.rl - Genbank #AA398976; qf47dll.x - Genbank #AI217144) corresponding to the 3' end of the human *vasa* gene were identified. These appeared to be bona fide human *vasa* EST's based on the extent of sequence conservation and the fact that all were derived from a testis source, consistent with germ cell origin. A 28-mer oligonucleotide corresponding to the 3' untranslated region was designed nad synthesized (5'-CTC TGC ATC AAA ACC ACA GAC TTG AAG G-3', SEQ ID NO:8). This oligonucleotide was then used in a 5' RACE (rapid amplification of cDNA ends) reaction to obtain a human *vasa* cDNA. Marathon-Ready human testis cDNA was obtained from CLONTECH and 5' RACE was performed as per the manufacturer's instructions, except that LA Taq (TAKARA) was utilized to minimize errors during PCR. A single main reaction product consisting of a 2.2 kilobase DNA fragment was obtained, purified by agarose gel electrophoresis, and subcloned into a commercially available cloning vector (pCR2.1, INVITROGEN) to yield plasmid pVAS3. The nucleotide sequence of the human *vasa* cDNA was obtained by double-stranded DNA sequencing; both strands were sequenced. The sequence of the human *vasa* protein (SEQ ID NO:2) was determined by conceptual translation of the cDNA sequence (SEQ ID NO:1).

Example 2: Preparation and Testing of Antibodies Selective for a *Vasa* Protein.

Regions likely to be antigenic were identified by an algorithm for antigencity (University of Wisconsin Genetics Computer Group Software). Two regions were selected for peptide synthesis: (1) CEDNPTRNRGFSKRGGYRDGNNSEASGPYR, SEQ ID NO:9 (amino acids 117-146 of SEQ ID NO:2); and (2) VDTRKGKSTLNTAGFSSSRAPNPVDDESW, SEQ ID NO:10 (amino acids 695-723 of

SEQ ID NO:2). As an alternative approach to generating antigen, the N-terminus of the *vasa* polypeptide (amino acids 1-318 of SEQ ID NO:2), which is believed to be highly antigenic, was expressed in *E. coli*. PCR was performed on pVAS3 with the primers 5'-AAG TCA CCA TGG GGG ATG AA-3', SEQ ID NO:11 (designed around a naturally occuring *NcoI* site) and 5'-TTA AGA TCT TTT TTG CAC AGG AGT AAG C-3', SEQ ID NO:12 (which contains an engineering *BglII* site). The product was digested with the restriction enzymes *NcoI* and *BglII* and ligated in the expression vector PQE-60 (QIAGEN) linearized with *NcoI* and *BglII* (directional cloning). This results in an in-frame expression construct where the first amino acid of the expressed protein corresponds to the native methionine and the C-terminus is fused to six consecutive histidines encoded by the vector (for protein purification). Following expression in *E. coli* and affinity purification of the protein according to the manufacturer's instructions, the expressed protein was used to immunize animals according to methods known to one of ordinary skill in the art.

Anti-*vasa* antibodies were prepared and tested for selectivity for the *vasa* protein using conventional immunohisto- and immunocyto- chemistry methods adapted to include the *vasa* protein (or fragment thereof) as a positive control and a non-*vasa* protein (lacking sequence homology to *vasa* protein) as a negative control. Additional controls (e.g., negative controls containing nongonadal tissue, tumors of non-germ cell origin), can be included in the assay to establish the specificity of the antibodies for detecting tumors of germ cell origin in a complex sample.

The preferred assays for tissue samples are performed on formalin-fixed, paraffin-embedded tissue following antigen retrieval as routinely done on clinical material in a hospital pathology laboratory setting. Exemplary samples (including positive and negative controls) include: normal testes and ovaries, other normal tissues, a range of major tumor types (to evaluate specificity) and germ cell tumors of various histologic subtypes and anatomic locations (to establish clinical utility for detecting these tumors). Preliminary results confirmed the predicted cytoplasmic immunoreactivity in germ cells with an appropriate negative result obtained for the preimmune control.

In a further example, the specificity of anti-human *vasa* antibodies for testicular germ cell tumors was also confirmed using immunohistochemical stains. Granulosa cell tumor (non-germ cell tumor of the ovary -negative control) showed no positive stain. By contrast, classic seminoma (testis) showed strong positive reaction (appearing as brown stain) only in tumor cells and not in the surrounding stroma. Permatocytic seminoma (testis) also showed strong positive reaction in in all tumor cells. For this particular experiment, sections were

incubated with a 1:1000 dilution of affinity-purified polyclonal anti-*vasa* antibody. Detection was performed as described elsewhere herein using diaminobenzidine as the chromogen. Slides were counterstained with hematoxylin.

Example 3: Detection of A *Vasa* Protein By Immunoassay.

Vasa protein (peptide and/or polypeptide) detection is carried out by the indirect enzyme-linked immunosorbent assay (ELISA). Such assays are well known in the art. Briefly, ninety six-well microtiter plates (Dynatech Laboratories, Alexandria, VA) are coated with, for example, serum, germ cell tumor culture supernatants, germ cell tumor lysates/extracts, etc., and left overnight at 4°C. The plates were then washed 3 times with PBS containing 0.05% Tween-20 and blocked with 5% dry milk for 2 hrs at 37 °C. The plates are washed and incubated again overnight at 4°C with polyclonal *vasa* antisera or *vasa* monoclonal antibodies as described earlier (diluted 1:100 -1:10^7 in PBS with 1% dry milk). The plates are washed, and 100µl of peroxidase-conjugated goat anti-mouse (or anti-*animal* used to generate antibodies/sera) IgG (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) at 1:2,000 dilution in PBS with 0.05% Tween-20 and 1% dry milk, are added. The plates are incubated for 2 hrs at 37 °C. The plates are then washed 3 times with PBS containing 0.05% Tween-20. Then, 100µl of O-phenylenediamine (0.4mg/ml in citrate phosphate buffer containing 0.015% hydrogen peroxide; Sigma Chemical Co. St Louis, MO), are added to each well, and the reaction is stopped by the addition of 2.5N HCl. After 15 minutes the optical density at 492nm (OD_{492}) is measured using a EL308 ELISA reader (Bio-Tek Instruments, Winooski,VT). The titer expressed in OD_{492} unit is calculated as a multiple of the dilution in the linear portion of a standard plot.

Example 4: Detection of *Vasa* Nucleic Acid by PCR amplification.

An exemplary procedure for PCR amplification to detect a tumor marker is provided in U.S. Patent No. 5,688,648, entitled, "Methods of detecting micrometastasis of prostate cancer", issued to C. Croce, et al. It is to be understood that the exemplary procedures described therein can be modified by one of ordinary skill in the art using no more that routine experimentation to detect the *vasa* molecule in any type of sample material using the novel reagents disclosed herein.

Vasa protein expression is detected by determining whether mRNA for the *vasa* protein is present in a sample. The preferred procedure for detecting mRNA for *vasa* protein is by PCR amplification.

Synthetic oligonucleotides: 5'-TGC ATC AAA ACC ACA GAC TTG-3', SEQ ID NO:13, and 5'-AAT GCC ATC AAA GGA ACA GC-3', SEQ ID NO:14, were designed using the Primer3 program (Whitehead Institute for Biomedical Research) for an RT-PCR assay of *vasa* mRNA expression. RT-PCR was performed using the Superscript One-Step PCR kit (GIBCO) per the manufacturer's instructions. When RT-PCR was performed on total human testis RNA, a single product of expected size (804 bp) was consistently detected. However, no products were detected with RNA from extragonadal sources (e.g., liver), confirming the specificity of the assay.

Example 5: Detection of *vasa* molecules in Tumor Test Samples

An exemplary procedure for detecting a tumor marker in a breast tumor tissue sample or extract is provided in U.S. Patent No. 5,723,302, entitled, "Detection of prostrate-specific antigen in breast tumors", issued to E. Diamandis. An exemplary procedure for detecting a tumor marker in a blood sample is provided in U.S. Patent No. 5,935,775, entitled, "Whole blood analysis of prostate specific antigen spotted on a solid support", issued to G. Savjani. It is to be understood that these exemplary procedures can be modified by one of ordinary skill in the art using no more that routine experimentation to detect the *vasa* molecule in any type of sample using the novel reagents disclosed herein.

Sample Preparation.

(1) Solid tissue/tumor samples: A tumor sample or intact tissue is prepared in accordance with standard histological procedures and analyzed by, e.g., *in situ* immunohistochemistry, enzyme immunoassay (to detect *vasa* protein), or by, e.g., nucleic acid enrichment (to detect a *vasa* nucleic acid). Immunoassays are performed as described above (Example 3) using the anti-*vasa* protein antibodies that selectively bind to the *vasa* protein or peptides. Nucleic acid assays are performed as described above (Example 4) for detection of a *vasa* nucleic acid, e.g., by PCR amplification. Total RNA or mRNA is isolated from the tumor samples and cDNA synthesized by reverse transcription. PCR amplification of cDNA is accomplished using *vasa* specific primers also as described above. A probe is used to detect cDNA for *vasa*. Other methods for detecting an RNA for *vasa* may also be used, such as, the northern blot technique.

(2) Blood samples: Blood samples can be used directly in liquid form, dried onto a solid support (see, e.g., U.S. 5,935,775), or can be further processed in accordance with standard procedures known to those of ordinary skill in the art for use in the assays of the invention. For example, the sera can be separated into fractions prior to analysis for the presence of a *vasa* molecule.

Analysis.

It is anticipated that the presence of a *vasa* molecule in a tumor sample of a subject is diagnostic of the presence of a tumor of germ cell origin in the subject. To validate the detection methods disclosed herein for predicting a tumor of germ cell origin, samples of known pathology are assayed to determine the presence or absence of a *vasa* molecule and the presence or absence of the *vasa* molecule is correlated to the presence or absence of a tumor of germ cell origin in known samples of positive controls (tumor present) and negative controls (tumor absent), respectively. Additional internal assay controls can be used to verify assay reproducibility (e.g., analysis of known sample components, such as other known protein or other known nucleic acid components in the tissue from which the sample is derived, e.g., β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP -glial fibrillary acidic protein, synaptophysin, chromogranin, cytokeratin isoforms such as 7 and 20, anti-keratin markers such as AE1/AE3 and CAM5.2, etc.). A statistically significant, positive correlation is established between the presence of a *vasa* molecule and the presence of a tumor of germ cell origin. A statistically significnt positive correlation also is established between the relative amount of the *vasa* molecule present in the sample (e.g., of gonadal origin) of a subject, and the stage and/or size of the subject's tumor, as well as the degree of metastasis in the subject. This correlation is used to evaluate patient response to treatment modalities, monitor tumor regression and/or remission, and/or to predict patient survival.

Table I. Sequences with partial homologies to human *vasa*

Sequences with GenBank accession numbers:
S75275, D14859, AB005147, Y12007, AF046043, Z81449.1, X81823, P09052, Q64060, Q61496, Q62167, O00571, P24346, P16381, O15523, AL042306, AA399611, AA398976, AA383535, AI217144, AI953070, AI025074, AI654417, AI337133, AA758412, AI969018, AA400066, AA862553, AA401568, AA316798, T85890, T82153.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

All references disclosed herein are incorporated by reference in their entirety.

What is claimed is presented below and is followed by a Sequence Listing.

We claim:

Claims

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleotide sequence set forth as SEQ ID NO:1 and which code for a human *vasa* polypeptide,

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, or

(c) complements of (a) or (b).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises the nucleotide sequence set forth as SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule consists of the nucleotide sequence set forth as SEQ ID NO:15 or a fragment thereof.

4. An isolated nucleic acid molecule selected from the group consisting of

(a) unique fragments of a nucleotide sequence set forth as SEQ ID NO:1,

(b) complements of (a),

provided that a unique fragment of (a) includes a sequence of contiguous nucleotides which is not identical to any sequence selected from the sequence group consisting of

(1) sequences having the database accession numbers of Table I, or sequences encoding a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, and SEQ ID NO:7,

(2) complements of (1), and

(3) fragments of (1) and (2).

5. The isolated nucleic acid molecule of claim 4, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

(1) at least two contiguous nucleotides nonidentical to the sequence group,

(2) at least three contiguous nucleotides nonidentical to the sequence group,

(3) at least four contiguous nucleotides nonidentical to the sequence group,

(4) at least five contiguous nucleotides nonidentical to the sequence group,

(5) at least six contiguous nucleotides nonidentical to the sequence group,

(6) at least seven contiguous nucleotides nonidentical to the sequence group.

6. The isolated nucleic acid molecule of claim 4, wherein the fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

7. The isolated nucleic acid molecule of claim 4, wherein the molecule encodes a polypeptide which is immunogenic.

8. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, or 7 operably linked to a promoter.

9. An expression vector comprising the isolated nucleic acid molecule of claim 4 operably linked to a promoter.

10. A host cell transformed or transfected with the expression vector of claim 8.

11. A host cell transformed or transfected with the expression vector of claim 9.

12. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, or 4, wherein the polypeptide, or fragment of the polypeptide, has germ cell specific expression.

13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2.

14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids 1-724 of SEQ ID NO:2.

15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, or 4, wherein the polypeptide, or fragment of the polypeptide, is immunogenic.

16. The fragment of claim 15, wherein the fragment, or portion of the fragment, binds to a human antibody.

17. An isolated binding polypeptide which binds selectively a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3 or 4.

18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids of SEQ ID NO:2.

19. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids of SEQ ID NO:9 or SEQ ID NO:10.

20. The isolated binding polypeptide of any one of claims 18 or 19, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a $F(ab)_2$ fragment or a fragment including a CDR3 region selective for the polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:9, and SEQ ID NO:10.

21. An isolated polypeptide comprising a fragment of the polypeptide of claim 12 having germ cell specific expression, provided that the fragment excludes a sequence of contiguous amino acids selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, and SEQ ID NO:7, or is encoded by an isolated nucleic acid having a nucleotide sequence with a GenBank database accession number as described in Table I.

22. A kit, comprising a package containing:

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.

23. The kit of claim 22, wherein the control is a predetermined value for comparing to the measured value.

24. The kit of claim 23, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.

25. A method for determining the level of a *vasa* molecule expression in a subject, comprising:

a) obtaining a test sample from a subject,

b) measuring the expression of a *vasa* molecule in the test sample,

c) comparing the measured expression of the *vasa* molecule to a control.

26. The method of claim 25, wherein the expression of a *vasa* molecule in (b) is *vasa* mRNA expression.

27. The method of claim 25, wherein the expression of a *vasa* molecule in (b) is *vasa* polypeptide expression.

28. The method of claim 25, wherein the test sample is tissue.

29. The method of claim 25, wherein the test sample is a biological fluid.

30. The method of claim 25, wherein the test sample is a fine needle aspirate.

31. The method of claim 26, wherein *vasa* mRNA expression is measured using the Polymerase Chain Reaction (PCR).

32. The method of claim 26, wherein *vasa* mRNA expression is measured using northern blotting.

33. The method of claim 27, wherein *vasa* polypeptide expression is measured using a monoclonal antibody to a *vasa* polypeptide.

34. The method of claim 27, wherein *vasa* polypeptide expression is measured using polyclonal antisera to a *vasa* polypeptide.

35. A method of detecting a tumor of germ cell origin in an extragonadal sample, the method comprising:

detecting *vasa* expression in an extragonadal test sample obtained from a subject, wherein *vasa* expression in the extragonadal test sample is indicative of a tumor of germ cell origin in the subject.

36. The method of claim 35, wherein *vasa* expression is *vasa* mRNA expression.

37. The method of claim 35, wherein *vasa* expression is *vasa* polypeptide expression.

38. The method of claim 35, wherein the extragonadal test sample is tissue.

39. The method of claim 35, wherein the extragonadal test sample is a biological fluid.

40. The method of claim 35, wherein the test sample is a fine needle aspirate.

41. The method of claim 36, wherein *vasa* mRNA expression is measured using the Polymerase Chain Reaction.

42. The method of claim 36, wherein *vasa* mRNA expression is measured using northern blotting.

43. The method of claim 37, wherein *vasa* polypeptide expression is measured using a monoclonal antibody to a *vasa* polypeptide.

44. The method of claim 37, wherein *vasa* polypeptide expression is measured using polyclonal antisera to a *vasa* polypeptide.

45. The method of claim 35, wherein the subject has not previously been diagnosed as having a tumor of germ cell origin or a predisposition thereto.

46. The method of claim 35, wherein the subject has a clinical diagnosis of a tumor of germ cell origin and the method is to confirm the clinical diagnosis, monitor a remission of the tumor, or stage the tumor.

47. A method of detecting a tumor of germ cell origin, the method comprising:

detecting *vasa* overexpression in a test sample obtained from a subject, wherein *vasa* overexpression in the test sample as compared to a control is indicative of a tumor of germ cell origin in the subject.

48. The method of claim 47, wherein *vasa* overexpression is *vasa* mRNA overexpression.

49. The method of claim 47, wherein *vasa* overexpression is *vasa* polypeptide overexpression.

50. The method of claim 47, wherein the test sample is tissue.

51. The method of claim 47, wherein the test sample is a biological fluid.

52. The method of claim 47, wherein the test sample is a fine needle aspirate.

53. The method of claim 48, wherein *vasa* mRNA overexpression is measured using the Polymerase Chain Reaction (PCR).

54. The method of claim 48, wherein *vasa* mRNA overexpression is measured using northern blotting.

55. The method of claim 49, wherein *vasa* polypeptide overexpression is measured using a monoclonal antibody to a *vasa* polypeptide.

56. The method of claim 49, wherein *vasa* polypeptide overexpression is measured using polyclonal antisera to a *vasa* polypeptide.

57. The method of claim 47, wherein the tumor is selected from the group consisting of a testicular tumor, an ovarian tumor, and a tumor of an extragonadal tissue.

58. The method of claim 47, wherein the tumor is a testicular tumor.

59. The method of claim 47, wherein the tumor is an ovarian tumor.

60. The method of claim 47, wherein the tumor is a tumor of an extragonadal tissue.

61. The method of claim 47, wherein the tumor is a seminoma.

62. The method of claim 47, further comprising detecting expression of a tumor-specific agent other than a *vasa* molecule in the test sample.

63. The method of claim 62, wherein the tumor-specific agent other than a *vasa* molecule is selected from the group consisting of β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP, synaptophysin, chromogranin, cytokeratin isoforms, and anti-keratin markers.

64. The method of claim 47, wherein the subject has not previously been diagnosed as having a tumor of germ cell origin or a predisposition thereto.

65. The method of claim 47, wherein the subject has a clinical diagnosis of a tumor of germ cell origin and the method is to confirm the clinical diagnosis, monitor a remission of the tumor, or stage the tumor.

66. A method of subtyping tumors of germ cell origin, comprising:

detecting *vasa* expression in a test sample of a known or suspected tumor of germ cell origin obtained from a subject, wherein *vasa* overexpression in the test sample as compared to a control is indicative of a seminoma in the subject, or wherein absence of *vasa* expression in the test sample as compared to a control is indicative of a nonseminoma in the subject.

67. The method of claim 66, wherein *vasa* expression is *vasa* mRNA expression.

68. The method of claim 66, wherein *vasa* expression is *vasa* polypeptide expression.

69. The method of claim 66, wherein the test sample is tissue.

70. The method of claim 66, wherein the test sample is a biological fluid.

71. The method of claim 66, wherein the test sample is a fine needle aspirate.

72. The method of claim 67, wherein *vasa* mRNA expression is measured using the Polymerase Chain Reaction (PCR).

73. The method of claim 67, wherein *vasa* mRNA expression is measured using northern blotting.

74. The method of claim 68, wherein *vasa* polypeptide expression is measured using a monoclonal antibody to a *vasa* polypeptide.

75. The method of claim 68, wherein *vasa* polypeptide expression is measured using polyclonal antisera to a *vasa* polypeptide.

76. The method of claim 66, wherein the tumor is selected from the group consisting of a testicular tumor, an ovarian tumor, and a tumor of an extragonadal tissue.

77. The method of claim 66, wherein the tumor is a testicular tumor.

78. The method of claim 66, wherein the tumor is an ovarian tumor.

79. The method of claim 66, wherein the tumor is a tumor of an extragonadal tissue.

80. The method of claim 66, wherein the nonseminoma is selected from the group consisting of an embryonal carcinoma, a teratoma, a choriocarcinoma, a yolk sac tumor, or combinations of the foregoing.

81. The method of claim 66, wherein the subject has not previously been diagnosed as having a tumor of germ cell origin or a predisposition thereto.

82. The method of claim 66, wherein the subject has a clinical diagnosis of a tumor of germ cell origin and the method is to confirm the clinical diagnosis, monitor a remission of the tumor, or stage the tumor.

83. The method of claim 66, wherein the subject has a clinical diagnosis of a tumor of mixed histologic appearance.

84. The method of claim 66, further comprising detecting expression of a tumor-specific agent other than a *vasa* molecule in the test sample.

85. The method of claim 84, wherein the tumor-specific agent other than a *vasa* molecule is selected from the group consisting of β-hCG, α-fetoprotein, placental-type alkaline phosphatase, prostate specific antigen, carcinoembryonic antigen, inhibin, epithelial membrane antigen, desmin, vimentin, GFAP, synaptophysin, chromogranin, cytokeratin isoforms, and anti-keratin markers.

86. A method of distinguishing a tumor of germ cell origin from a non-germ cell tumor, the method comprising:

detecting expression of a *vasa* molecule in a test sample, wherein expression of the *vasa* molecule is indicative of a tumor of a germ cell origin and absence of expression of the *vasa* molecule is indicative of a non-germ cell tumor.

87. The method of claim 86, wherein the non-germ cell tumor resembles histologically a tumor of germ cell origin.

88. The method of claim 86, wherein the non-germ cell tumor is selected from the group consisting of a clear cell carcinoma of the ovary, a mediastinal thymoma, and a testicular lymphoma.

Abstract

The invention pertains to nucleic acids encoding a human *vasa* protein, including fragments and biologically functional variants thereof. The invention also pertains to therapeutics and diagnostics involving the foregoing proteins and genes and agents that bind the foregoing proteins and genes.

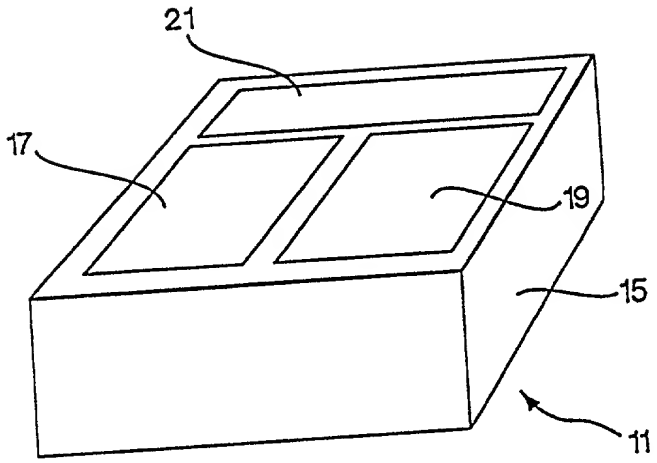


Fig. 1

SEQUENCE LISTING

<110> Castrillon, Diego H.

<120> COMPOSITIONS AND METHODS FOR THE IMPROVED DIAGNOSIS OF GERM CELL TUMORS

<130> B0801/7195/ERP/KA

<150> U.S. 60/166,394
<151> 1999-11-18

<160> 15

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2224
<212> DNA
<213> Homo Sapiens

<400> 1

acttgaagtc accatggggg atgaagattg ggaagcagaa atcaaccctc atatgtcttc 60
ctatgttccc atatttgaga aggataggta ttctggagaa aatggagaca attttaacag 120
gactccagct tcatcatcag aaatggatga tggaccttct cgaagagatc atttcatgaa 180
aagtggattt gcctctgggc ggaattttgg aaacagagat gctggtgagt gtaataagcg 240
agataataca tccacaatgg gtggttttgg agttggaaag agttttggaa acagaggttt 300
ttcaaacagc aggtttgaag atggtgatag ctctggtttc tggagagagt ctagtaatga 360
ctgcgaagat aatccaacac ggaacagagg gttttccaag agaggcggct atcgagatgg 420
aaataattca gaagcttcag ggccatacag aagaggtgga agaggtagtt tccgaggttg 480
ccgtggagga tttggtctag gaagtccaaa taatgactta gacccagacg aatgtatgca 540
gcgcactggt ggccttttg gttctagaag accagtatta agtggcacag gtaatggtga 600
tacttctcaa agcagaagtg gcagtggaag tgaacgaggt ggttacaaag gtttaaatga 660
agaagtaata acaggctctg gaaagaattc ttggaagtca gaagcagaag gaggagaaag 720
tagtgatact caaggaccaa aagtgaccta catacccccc cctccacctg aggatgagga 780
ctccatcttt gcacattatc agacaggcat aaacttcgac aaatacgaca ctattcttgt 840
ggaagtgtct ggacatgatg caccaccagc aattctgact tttgaagaag ctaatctctg 900
tcagacactg aataacaaca ttgctaaagc tggttatact aagcttactc ctgtgcaaaa 960
atacagtatt cctatcatac ttgcaggacg agatttgatg gcttgtgctc aaacagggtc 1020
tgggaagact gcggcttttc tcctaccaat tttggctcat atgatgcatg atggaataac 1080
tgccagtcgt tttaaagagt tgcaggaacc agagtgtatt attgtagcac caactcgaga 1140
attggtcaac cagatttatt tggaagccag aaaattttct tttgggactt gtgtaagagc 1200
tgttgttata tatgggggaa cccagctggg acattcaatt cgacaaatag tacaaggctg 1260
taatatatta tgtgctactc ctggaagact gatggatatc ataggcaaag aaaagattgg 1320
tctcaaacag atcaaatact tagttttgga tgaagctgat cgcatgttgg atatgggttt 1380
tggtccagaa atgaagaagt taatttcttg cccaggaatg ccatcaaagg aacagcgcca 1440
aacccttatg ttcagtgcaa cttttccaga ggaaattcaa aggttggctg cagagttttt 1500
aaagtcaaat tatctgtttg ttgctgttgg acaagtgggt ggagcatgta gagatgttca 1560
gcagaccgtt ctccaagttg gccagttctc aaaaagagaa aagctcgttg aaattctgcg 1620
aaacataggg gatgaaagaa ctatggtctt tgttgaaact aagaaaaaag cagattttac 1680
tgcaactttt ctttgtcaag aaaaaatatc aactacaagt atccatggtg atcgggaaca 1740
gagagagcgg gagcaagctc ttggagattt tcgctttgga aagtgcccag ttcttgttgc 1800
tacttcagta gctgccagag ggctggatat tgaaaatgtg caacatgtta tcaattttga 1860
tcttccttct accattgatg aatatgttca tcgaattggg cgtactggtc gttgtgggaa 1920
tactggcaga gcaatttcct tttttgatct tgaatcggat aaccatttag cacagcctct 1980
agtaaaagta ttgacagatg ctcaacagga tgttcctgca tggttggaag aaattgcctt 2040
tagtacatac attcctggct tcagtggtag tacaagagga aacgtgtttg catcagttga 2100

taccagaaag ggcaagagca ctttgaacac agctgggttt tcttcttcac gagctcccaa 2160
tccagtagat gatgagtcat gggattaaag ccaaaacatc cttcaagtct gtggttttga 2220
tgca 2224

<210> 2
<211> 724
<212> PRT
<213> Homo Sapiens

<400> 2

```
Met Gly Asp Glu Asp Trp Glu Ala Glu Ile Asn Pro His Met Ser Ser
 1               5                  10                  15
Tyr Val Pro Ile Phe Glu Lys Asp Arg Tyr Ser Gly Glu Asn Gly Asp
            20                  25                  30
Asn Phe Asn Arg Thr Pro Ala Ser Ser Ser Glu Met Asp Asp Gly Pro
        35                  40                  45
Ser Arg Arg Asp His Phe Met Lys Ser Gly Phe Ala Ser Gly Arg Asn
    50                  55                  60
Phe Gly Asn Arg Asp Ala Gly Glu Cys Asn Lys Arg Asp Asn Thr Ser
65                  70                  75                  80
Thr Met Gly Gly Phe Gly Val Gly Lys Ser Phe Gly Asn Arg Gly Phe
                85                  90                  95
Ser Asn Ser Arg Phe Glu Asp Gly Asp Ser Ser Gly Phe Trp Arg Glu
            100                 105                 110
Ser Ser Asn Asp Cys Glu Asp Asn Pro Thr Arg Asn Arg Gly Phe Ser
        115                 120                 125
Lys Arg Gly Gly Tyr Arg Asp Gly Asn Asn Ser Glu Ala Ser Gly Pro
    130                 135                 140
Tyr Arg Arg Gly Gly Arg Gly Ser Phe Arg Gly Cys Arg Gly Gly Phe
145                 150                 155                 160
Gly Leu Gly Ser Pro Asn Asn Asp Leu Asp Pro Asp Glu Cys Met Gln
                165                 170                 175
Arg Thr Gly Gly Leu Phe Gly Ser Arg Arg Pro Val Leu Ser Gly Thr
            180                 185                 190
Gly Asn Gly Asp Thr Ser Gln Ser Arg Ser Gly Ser Gly Ser Glu Arg
        195                 200                 205
Gly Gly Tyr Lys Gly Leu Asn Glu Glu Val Ile Thr Gly Ser Gly Lys
    210                 215                 220
Asn Ser Trp Lys Ser Glu Ala Glu Gly Gly Glu Ser Ser Asp Thr Gln
225                 230                 235                 240
Gly Pro Lys Val Thr Tyr Ile Pro Pro Pro Pro Pro Glu Asp Glu Asp
                245                 250                 255
Ser Ile Phe Ala His Tyr Gln Thr Gly Ile Asn Phe Asp Lys Tyr Asp
            260                 265                 270
Thr Ile Leu Val Glu Val Ser Gly His Asp Ala Pro Pro Ala Ile Leu
        275                 280                 285
Thr Phe Glu Glu Ala Asn Leu Cys Gln Thr Leu Asn Asn Asn Ile Ala
    290                 295                 300
Lys Ala Gly Tyr Thr Lys Leu Thr Pro Val Gln Lys Tyr Ser Ile Pro
305                 310                 315                 320
Ile Ile Leu Ala Gly Arg Asp Leu Met Ala Cys Ala Gln Thr Gly Ser
                325                 330                 335
Gly Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu Ala His Met Met His
            340                 345                 350
Asp Gly Ile Thr Ala Ser Arg Phe Lys Glu Leu Gln Glu Pro Glu Cys
        355                 360                 365
Ile Ile Val Ala Pro Thr Arg Glu Leu Val Asn Gln Ile Tyr Leu Glu
    370                 375                 380
```

```
Ala Arg Lys Phe Ser Phe Gly Thr Cys Val Arg Ala Val Val Ile Tyr
385                 390                 395                 400
Gly Gly Thr Gln Leu Gly His Ser Ile Arg Gln Ile Val Gln Gly Cys
                405                 410                 415
Asn Ile Leu Cys Ala Thr Pro Gly Arg Leu Met Asp Ile Ile Gly Lys
            420                 425                 430
Glu Lys Ile Gly Leu Lys Gln Ile Lys Tyr Leu Val Leu Asp Glu Ala
        435                 440                 445
Asp Arg Met Leu Asp Met Gly Phe Gly Pro Glu Met Lys Lys Leu Ile
    450                 455                 460
Ser Cys Pro Gly Met Pro Ser Lys Glu Gln Arg Gln Thr Leu Met Phe
465                 470                 475                 480
Ser Ala Thr Phe Pro Glu Glu Ile Gln Arg Leu Ala Ala Glu Phe Leu
                485                 490                 495
Lys Ser Asn Tyr Leu Phe Val Ala Val Gly Gln Val Gly Gly Ala Cys
            500                 505                 510
Arg Asp Val Gln Gln Thr Val Leu Gln Val Gly Gln Phe Ser Lys Arg
        515                 520                 525
Glu Lys Leu Val Glu Ile Leu Arg Asn Ile Gly Asp Glu Arg Thr Met
    530                 535                 540
Val Phe Val Glu Thr Lys Lys Lys Ala Asp Phe Thr Ala Thr Phe Leu
545                 550                 555                 560
Cys Gln Glu Lys Ile Ser Thr Thr Ser Ile His Gly Asp Arg Glu Gln
                565                 570                 575
Arg Glu Arg Glu Gln Ala Leu Gly Asp Phe Arg Phe Gly Lys Cys Pro
            580                 585                 590
Val Leu Val Ala Thr Ser Val Ala Ala Arg Gly Leu Asp Ile Glu Asn
        595                 600                 605
Val Gln His Val Ile Asn Phe Asp Leu Pro Ser Thr Ile Asp Glu Tyr
    610                 615                 620
Val His Arg Ile Gly Arg Thr Gly Arg Cys Gly Asn Thr Gly Arg Ala
625                 630                 635                 640
Ile Ser Phe Phe Asp Leu Glu Ser Asp Asn His Leu Ala Gln Pro Leu
                645                 650                 655
Val Lys Val Leu Thr Asp Ala Gln Gln Asp Val Pro Ala Trp Leu Glu
            660                 665                 670
Glu Ile Ala Phe Ser Thr Tyr Ile Pro Gly Phe Ser Gly Ser Thr Arg
        675                 680                 685
Gly Asn Val Phe Ala Ser Val Asp Thr Arg Lys Gly Lys Ser Thr Leu
    690                 695                 700
Asn Thr Ala Gly Phe Ser Ser Ser Arg Ala Pro Asn Pro Val Asp Asp
705                 710                 715                 720
Glu Ser Trp Asp
```

<210> 3
<211> 722
<212> PRT
<213> Mus Musculus

<400> 3

```
Met Gly Asp Glu Asp Trp Glu Ala Glu Ile Leu Lys Pro His Val Ser
 1               5                  10                  15
Ser Tyr Val Pro Val Phe Glu Lys Asp Lys Tyr Ser Ser Gly Ala Asn
            20                  25                  30
Gly Asp Thr Phe Asn Arg Thr Ser Ala Ser Ser Glu Met Glu Asp Gly
        35                  40                  45
Pro Ser Gly Arg Asp Asp Phe Met Arg Ser Gly Phe Pro Ser Gly Arg
```

```
    50                  55                  60
Ser Leu Gly Ser Arg Asp Ile Gly Glu Ser Ser Lys Lys Glu Asn Thr
65                  70                  75                  80
Ser Thr Thr Gly Gly Phe Gly Arg Gly Lys Gly Phe Gly Asn Arg Gly
                85                  90                  95
Phe Leu Asn Asn Lys Phe Glu Glu Gly Asp Ser Ser Gly Phe Trp Lys
            100                 105                 110
Glu Ser Asn Asn Asp Cys Glu Asp Asn Gln Thr Arg Ser Arg Gly Phe
        115                 120                 125
Ser Lys Arg Gly Gly Cys Gln Asp Gly Asn Asp Ser Glu Ala Ser Gly
    130                 135                 140
Pro Phe Arg Arg Gly Gly Arg Gly Ser Phe Arg Gly Cys Arg Gly Gly
145                 150                 155                 160
Phe Gly Leu Gly Arg Pro Asn Ser Glu Ser Asp Gln Asp Gln Gly Thr
                165                 170                 175
Gln Cys Gly Gly Gly Phe Leu Val Leu Gly Lys Pro Ala Ala Ser Asp
            180                 185                 190
Ser Gly Asn Gly Asp Thr Tyr Gln Ser Arg Ser Gly Ser Gly Arg Gly
        195                 200                 205
Gly Tyr Lys Gly Leu Asn Glu Glu Val Val Thr Gly Ser Gly Lys Asn
    210                 215                 220
Ser Trp Lys Ser Glu Thr Glu Gly Gly Glu Ser Ser Asp Ser Gln Gly
225                 230                 235                 240
Pro Lys Val Thr Tyr Ile Pro Pro Pro Pro Pro Glu Asp Glu Asp Ser
                245                 250                 255
Ile Phe Ala His Tyr Gln Thr Gly Ile Asn Phe Asp Lys Tyr Asp Thr
            260                 265                 270
Ile Leu Val Glu Val Ser Gly His Asp Ala Pro Pro Ala Ile Leu Thr
        275                 280                 285
Phe Glu Glu Ala Asn Leu Cys Gln Thr Leu Asn Asn Asn Ile Arg Lys
    290                 295                 300
Ala Gly Tyr Thr Lys Leu Thr Pro Val Gln Lys Tyr Thr Ile Pro Ile
305                 310                 315                 320
Val Leu Ala Gly Arg Asp Leu Met Ala Cys Ala Gln Thr Gly Ser Gly
                325                 330                 335
Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu Ala His Met Met Arg Asp
            340                 345                 350
Gly Ile Thr Ala Ser Arg Phe Lys Glu Leu Gln Glu Pro Glu Cys Ile
        355                 360                 365
Ile Val Ala Pro Thr Arg Glu Leu Ile Asn Gln Ile Tyr Leu Glu Ala
    370                 375                 380
Arg Lys Phe Ser Phe Gly Thr Cys Val Ile Ser Val Val Ile Tyr Gly
385                 390                 395                 400
Gly Thr Gln Phe Gly His Ser Val Arg Gln Ile Val Gln Gly Cys Asn
                405                 410                 415
Ile Leu Cys Ala Thr Pro Gly Arg Leu Met Asp Ile Ile Gly Lys Glu
            420                 425                 430
Lys Ile Gly Leu Lys Gln Val Lys Tyr Leu Val Leu Asp Glu Ala Asp
        435                 440                 445
Ser Met Leu Asp Met Gly Phe Ala Pro Glu Ile Lys Lys Leu Ile Ser
    450                 455                 460
Cys Pro Gly Met Pro Ser Lys Glu Gln His Gln Thr Leu Leu Phe Ser
465                 470                 475                 480
Ala Thr Phe Pro Glu Glu Ile Gln Arg Leu Ala Gly Asp Phe Leu Lys
                485                 490                 495
Ser Asn Tyr Leu Phe Val Ala Val Gly Gln Val Gly Gly Ala Cys Arg
            500                 505                 510
Asp Val Gln Gln Thr Ile Leu Gln Val Gly Gln Tyr Gln Lys Glu Lys
```

```
        515                 520                 525
Ser Leu Leu Arg Phe Tyr Glu Asn Ile Gly Asp Glu Arg Thr Met Val
    530                 535                 540
Phe Val Glu Thr Lys Lys Lys Ala Asp Phe Ile Ala Thr Phe Leu Cys
545                 550                 555                 560
Gln Glu Lys Ile Ser Ser Thr Ser Ile His Gly Asp Arg Glu Gln Arg
                565                 570                 575
Glu Arg Glu Gln Ala Leu Gly Asp Phe Arg Cys Gly Lys Cys Pro Val
            580                 585                 590
Leu Val Ala Thr Ser Val Ala Ala Arg Gly Leu Asp Ile Glu Asn Val
        595                 600                 605
Gln His Val Ile Asn Phe Asp Leu Pro Ser Thr Ile Asp Glu Tyr Val
    610                 615                 620
His Arg Ile Gly Arg Thr Gly Arg Cys Gly Asn Thr Gly Arg Ala Ile
625                 630                 635                 640
Ser Phe Phe Asp Thr Asp Ser Asp Asn His Leu Ala Gln Pro Leu Val
                645                 650                 655
Lys Val Leu Ser Asp Ala Gln Gln Asp Val Pro Ala Trp Leu Glu Glu
            660                 665                 670
Ile Ala Phe Ser Thr Tyr Val Pro Pro Ser Phe Ser Ser Ser Thr Arg
        675                 680                 685
Gly Gly Ala Val Phe Ala Ser Val Asp Thr Arg Lys Asn Tyr Gln Gly
    690                 695                 700
Lys Ala His Val Glu Tyr Ser Gly Asp Phe Phe Phe Thr Ser Ser Gln
705                 710                 715                 720
Ser Ser
```

<210> 4
<211> 713
<212> PRT
<213> Rattus Norvegicus

<400> 4

```
Met Gly Asp Glu Asp Trp Glu Ala Glu Ile Leu Lys Pro His Val Ser
 1                5                  10                  15
Ser Tyr Val Pro Val Phe Glu Lys Asp Lys Tyr Ser Ser Gly Ala Asn
            20                  25                  30
Gly Asp Thr Phe Asn Arg Thr Ser Ala Ser Ser Ser Glu Met Glu Asp
        35                  40                  45
Gly Pro Ser Gly Arg Asp His Phe Met Arg Ser Gly Phe Ser Ser Gly
    50                  55                  60
Arg Asn Leu Gly Asn Arg Asp Ile Gly Glu Ser Ser Lys Arg Glu Thr
65                  70                  75                  80
Thr Ser Thr Thr Gly Gly Phe Gly Arg Gly Lys Gly Phe Gly Asn Arg
                85                  90                  95
Gly Phe Leu Asn Asn Lys Phe Glu Glu Gly Asp Ser Ser Gly Phe Trp
            100                 105                 110
Lys Glu Ser Thr Asn Asp Cys Glu Asp Thr Gln Thr Arg Ser Arg Gly
        115                 120                 125
Phe Ser Lys Arg Gly Gly Tyr Pro Asp Gly Asn Asp Ser Glu Ala Ser
    130                 135                 140
Gly Pro Phe Arg Arg Gly Gly Arg Asp Ser Glu Tyr Asp Gln Asp Gln
145                 150                 155                 160
Gly Ser Gln Arg Gly Gly Gly Leu Phe Gly Ser Arg Lys Pro Ala Ala
                165                 170                 175
Ser Asp Ser Gly Ser Gly Asp Thr Phe Gln Ser Arg Ser Gly Asn Ala
            180                 185                 190
```

```
Arg Gly Ala Tyr Lys Gly Leu Asn Glu Glu Val Val Thr Gly Ser Gly
        195                 200                 205
Lys Asn Ser Trp Lys Ser Glu Ala Glu Gly Gly Glu Ser Ser Asp Ile
    210                 215                 220
Gln Gly Pro Lys Val Thr Tyr Ile Pro Pro Pro Pro Pro Glu Asp Glu
225                 230                 235                 240
Asp Ser Ile Phe Ala His Tyr Gln Thr Gly Ile Asn Phe Asp Lys Tyr
                245                 250                 255
Asp Thr Ile Leu Val Glu Val Ser Gly His Asp Ala Pro Pro Ala Ile
            260                 265                 270
Leu Thr Phe Glu Glu Ala Asn Leu Cys Gln Thr Leu Asn Asn Asn Ile
        275                 280                 285
Ala Lys Ala Gly Tyr Thr Lys Leu Thr Pro Val Gln Lys Tyr Ser Ile
    290                 295                 300
Pro Ile Val Leu Ala Gly Arg Asp Leu Met Ala Cys Ala Gln Thr Gly
305                 310                 315                 320
Ser Gly Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu Ala His Met Met
                325                 330                 335
Arg Asp Gly Ile Thr Ala Ser Arg Phe Lys Glu Leu Gln Glu Pro Glu
            340                 345                 350
Cys Ile Ile Val Ala Pro Thr Arg Glu Leu Ile Asn Gln Ile Tyr Leu
        355                 360                 365
Glu Ala Arg Lys Phe Ser Phe Gly Thr Cys Val Arg Ala Val Val Ile
    370                 375                 380
Tyr Gly Gly Thr Gln Phe Gly His Ser Ile Arg Gln Ile Val Gln Gly
385                 390                 395                 400
Cys Asn Ile Leu Cys Ala Thr Pro Gly Arg Leu Met Asp Ile Ile Gly
                405                 410                 415
Lys Glu Lys Ile Gly Leu Lys Gln Val Lys Tyr Leu Val Leu Asp Glu
            420                 425                 430
Ala Asp Arg Met Leu Asp Met Gly Phe Gly Pro Glu Met Lys Lys Leu
        435                 440                 445
Ile Ser Cys Pro Gly Met Pro Ser Lys Glu Gln Arg Gln Thr Leu Leu
    450                 455                 460
Phe Ser Ala Thr Phe Pro Glu Glu Ile Gln Arg Leu Ala Gly Glu Phe
465                 470                 475                 480
Leu Lys Ser Asn Tyr Leu Phe Val Ala Val Gly Gln Val Gly Gly Ala
                485                 490                 495
Cys Arg Asp Val Gln Gln Ser Ile Leu Gln Val Gly Pro Val Phe Lys
            500                 505                 510
Lys Arg Lys Leu Val Glu Ile Leu Arg Asn Ile Gly Asp Glu Arg Pro
        515                 520                 525
Met Val Phe Val Glu Thr Lys Lys Lys Ala Asp Phe Ile Ala Thr Phe
    530                 535                 540
Leu Cys Gln Glu Lys Ile Ser Thr Thr Ser Ile His Gly Asp Arg Glu
545                 550                 555                 560
Gln Arg Glu Arg Glu Gln Ala Leu Gly Asp Phe Arg Cys Gly Lys Cys
                565                 570                 575
Pro Val Leu Val Ala Thr Ser Val Ala Ala Arg Gly Leu Asp Ile Glu
            580                 585                 590
Asn Val Gln His Val Ile Asn Phe Asn Leu Pro Ser Thr Ile Asp Glu
        595                 600                 605
Tyr Val His Arg Ile Gly Arg Thr Gly Arg Cys Gly Asn Thr Gly Arg
    610                 615                 620
Ala Ile Ser Phe Phe Asp Thr Glu Ser Asp Asn His Leu Ala Gln Pro
625                 630                 635                 640
Leu Val Lys Val Leu Ser Asp Ala Gln Gln Asp Val Pro Ala Trp Leu
                645                 650                 655
```

Glu Glu Ile Ala Phe Ser Ser Tyr Ala Pro Pro Ser Phe Ser Asn Ser
660 665 670
Thr Arg Gly Ala Val Phe Ala Ser Phe Asp Thr Arg Lys Asn Phe Gln
675 680 685
Gly Lys Asn Thr Leu Asn Thr Ala Gly Ile Ser Ser Ala Gln Ala Pro
690 695 700
Asn Pro Val Asp Asp Glu Ser Trp Asp
705 710

<210> 5
<211> 700
<212> PRT
<213> Xenopus Laevis

<400> 5
Met Glu Glu Asn Trp Asp Thr Glu Ile Glu Thr Glu Lys Pro Thr Tyr
1 5 10 15
Val Pro Asn Phe Ser Thr Leu Glu Thr Glu Asn Thr Asp Asn Tyr Ser
20 25 30
Ala Tyr Ser Asn Asn Asp Ile Asn Asn Gln Asn Tyr Asp Ser Glu Arg
35 40 45
Ser Phe Gly Asn Arg Gly Gly Tyr Arg Ser Glu Arg Ser Arg Pro Ser
50 55 60
Asn Phe Asn Arg Gly Ser Arg Thr Glu Arg Gly Arg Gly Arg Gly Phe
65 70 75 80
Gly Thr Asn Arg Asn Asp Asn Tyr Ser Ser Glu Arg Asp Val Phe Gly
85 90 95
Asp Asp Glu Arg Asp Gln Arg Arg Gly Phe Pro Gly Arg Gly Gly Tyr
100 105 110
Asn Gly Asn Glu Asp Gly Gln Lys Pro Asn Ala Phe Arg Gly Arg Gly
115 120 125
Gly Phe Arg Asn Glu Asn Glu Gln Arg Arg Gly Phe Gly Glu Arg Gly
130 135 140
Gly Phe Arg Ser Glu Asn Gly Gln Arg Asn Phe Asp Asn Arg Gly Asp
145 150 155 160
Phe Gly Asn Ser Gly Glu Glu Glu Asp Arg Pro Arg Ser Tyr Gly Arg
165 170 175
Gly Gly Phe Asn Asn Ser Asp Thr Gly Gly Arg Gly Arg Arg Gly Gly
180 185 190
Arg Gly Gly Gly Ser Gln Tyr Gly Gly Tyr Lys Gly Arg Asn Glu Glu
195 200 205
Val Gly Val Glu Ser Gly Lys Ser Gln Glu Glu Gly Asn Glu Lys Asp
210 215 220
Glu Lys Pro Lys Lys Val Thr Tyr Ile Pro Pro Pro Pro Pro Asp Gly
225 230 235 240
Glu Asp Asn Ile Phe Arg Gln Tyr Gln Ser Gly Ile Asn Phe Asp Lys
245 250 255
Tyr Asp Glu Ile Leu Val Asp Val Thr Gly Lys Asp Val Pro Pro Ala
260 265 270
Ile Leu Thr Phe Glu Glu Ala Asn Leu Cys Glu Thr Leu Arg Arg Asn
275 280 285
Val Ala Arg Ala Gly Tyr Val Lys Leu Thr Pro Val Gln Lys His Ser
290 295 300
Ile Pro Ile Ile Met Ala Gly Arg Asp Leu Met Ala Cys Ala Gln Thr
305 310 315 320
Gly Ser Gly Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu Ser Tyr Met
325 330 335
Met Asn Glu Gly Ile Thr Ala Ser Gln Tyr Leu Gln Leu Gln Glu Pro

340 345 350
Glu Ala Ile Ile Ile Ala Pro Thr Arg Glu Leu Ile Asn Gln Ile Tyr
355 360 365
Leu Asp Ala Arg Lys Phe Ser Tyr Gly Thr Cys Val Arg Pro Val Val
370 375 380
Val Tyr Gly Gly Ile Gln Pro Val His Ala Met Arg Asp Val Glu Lys
385 390 395 400
Gly Cys Asn Ile Leu Cys Ala Thr Pro Gly Arg Leu Leu Asp Ile Val
405 410 415
Ser Lys Glu Lys Ile Gly Leu Ser Lys Leu Arg Tyr Leu Val Leu Asp
420 425 430
Glu Ala Asp Arg Met Leu Asp Met Gly Phe Ala Pro Glu Ile Glu Lys
435 440 445
Leu Met Thr Lys Pro Gly Met Pro Thr Lys Glu Lys Arg Gln Thr Leu
450 455 460
Met Phe Ser Ala Thr Tyr Pro Glu Glu Ile Arg Arg Leu Ala Ser Asn
465 470 475 480
Tyr Leu Lys Ser Glu His Leu Phe Val Val Val Gly Leu Val Gly Gly
485 490 495
Ala Cys Ser Asp Val Ala Gln Thr Val Leu Glu Met Arg Glu Asn Gly
500 505 510
Lys Met Glu Lys Leu Leu Glu Ile Leu Lys Ser Ser Glu Lys Glu Arg
515 520 525
Thr Met Ile Phe Val Asn Thr Lys Lys Lys Ala Asp Phe Ile Ala Gly
530 535 540
Tyr Leu Cys Gln Glu Lys Phe Ser Ser Thr Ser Ile His Gly Asp Arg
545 550 555 560
Glu Gln Tyr Gln Arg Glu Ser Ala Leu Trp Asp Phe Arg Thr Gly Lys
565 570 575
Cys Thr Val Ile Val Cys Thr Ala Val Ala Ala Arg Gly Leu Asp Ile
580 585 590
Glu Asn Val Gln His Val Ile Asn Tyr Asp Val Pro Lys Glu Val Asp
595 600 605
Glu Tyr Val His Arg Ile Gly Arg Thr Gly Arg Cys Gly Asn Thr Gly
610 615 620
Lys Ala Thr Ser Phe Phe Asn Val Gln Asp Asp His Val Ile Ala Arg
625 630 635 640
Pro Leu Val Lys Ile Leu Thr Asp Ala His Gln Glu Val Pro Ala Trp
645 650 655
Leu Glu Glu Ile Ala Phe Gly Gly His Gly Ala Leu Asn Ser Phe Tyr
660 665 670
Ala Ala Asp Ser Met Gly Glu Gln Ala Gly Gly Asn Ala Val Thr Thr
675 680 685
Pro Ser Phe Ala Gln Glu Glu Glu Ala Ser Trp Asp
690 695 700

<210> 6
<211> 700
<212> PRT
<213> Danio Reio

<400> 6
Met Asp Asp Trp Glu Glu Asp Gln Ser Pro Val Val Ser Cys Ser Ser
1 5 10 15
Gly Phe Gly Gly Ala Gly Asn Asp Lys Ser Asn Ser Glu Gly Thr Glu
20 25 30
Gly Ser Ser Trp Lys Met Thr Gly Asp Ser Phe Arg Gly Arg Gly Gly
35 40 45

Arg Gly Gly Ser Arg Gly Gly Arg Gly Gly Phe Ser Gly Phe Lys Ser
50 55 60

Glu Ile Asp Glu Asn Gly Ser Asp Gly Gly Trp Asn Gly Gly Glu Ser
65 70 75 80

Arg Gly Arg Gly Arg Gly Gly Phe Arg Gly Gly Phe Arg Ser Gly Ser
85 90 95

Arg Asp Glu Asn Asp Glu Asn Gly Asn Asp Asp Gly Trp Lys Gly Gly
100 105 110

Glu Ser Arg Gly Arg Gly Arg Gly Gly Phe Gly Gly Gly Phe Arg Gly
115 120 125

Gly Phe Arg Asp Gly Gly Asn Glu Asp Thr Gly Arg Arg Gly Phe Gly
130 135 140

Arg Glu Asn Asn Glu Asn Gly Asn Asp Glu Gly Gly Glu Gly Arg Gly
145 150 155 160

Arg Gly Arg Gly Gly Phe Arg Gly Gly Phe Arg Asp Gly Gly Gly Asp
165 170 175

Glu Ser Gly Lys Arg Gly Phe Gly Arg Gly Gly Phe Arg Gly Arg Asn
180 185 190

Glu Glu Val Phe Ser Lys Val Thr Thr Ala Asp Lys Leu Asp Gln Glu
195 200 205

Gly Ser Glu Asn Ala Gly Pro Lys Val Val Tyr Val Pro Pro Pro Pro
210 215 220

Pro Glu Glu Glu Ser Ser Ile Phe Ser His Tyr Ala Thr Gly Ile Asn
225 230 235 240

Phe Asp Lys Tyr Asp Asp Ile Leu Val Asp Val Ser Gly Ser Asn Pro
245 250 255

Pro Lys Ala Ile Met Thr Phe Glu Glu Ala Gly Leu Cys Asp Ser Leu
260 265 270

Ser Lys Asn Val Ser Lys Ser Gly Tyr Val Lys Pro Thr Pro Val Gln
275 280 285

Lys His Gly Ile Pro Ile Ile Ser Ala Gly Arg Asp Leu Met Ala Cys
290 295 300

Ala Gln Thr Gly Ser Gly Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu
305 310 315 320

Gln Arg Phe Met Thr Asp Gly Val Ala Ala Ser Lys Phe Ser Glu Ile
325 330 335

Gln Glu Pro Glu Ala Ile Ile Val Ala Pro Thr Arg Glu Leu Ile Asn
340 345 350

Gln Ile Tyr Leu Glu Ala Arg Lys Phe Ala Tyr Gly Thr Cys Val Arg
355 360 365

Pro Val Val Val Tyr Gly Gly Ile Asn Thr Gly Tyr Thr Ile Arg Glu
370 375 380

Val Leu Lys Gly Cys Asn Val Leu Cys Ala Thr Pro Gly Arg Leu His
385 390 395 400

Asp Leu Ile Gly Arg Gly Lys Ile Gly Leu Ser Lys Val Arg Tyr Leu
405 410 415

Val Leu Asp Glu Ala Asp Arg Met Leu Asp Met Gly Phe Glu Pro Glu
420 425 430

Met Arg Lys Leu Val Ala Ser Pro Gly Met Pro Ser Lys Glu Lys Arg
435 440 445

Gln Thr Leu Met Phe Ser Ala Thr Tyr Pro Glu Asp Ile Gln Arg Met
450 455 460

Ala Ala Asp Phe Leu Lys Val Asp Tyr Ile Phe Leu Ala Val Gly Val
465 470 475 480

Val Gly Gly Ala Cys Ser Asp Val Glu Gln Thr Ile Val Gln Val Asp
485 490 495

Gln Tyr Ser Lys Arg Asp Gln Leu Leu Glu Leu Leu Arg Ala Thr Gly
500 505 510

Asn Glu Arg Thr Met Val Phe Val Glu Thr Lys Arg Ser Ala Asp Phe
515 520 525
Ile Ala Thr Phe Leu Cys Gln Glu Lys Ile Ser Thr Thr Ser Ile His
530 535 540
Gly Asp Arg Glu Gln Arg Glu Arg Glu Lys Ala Leu Ser Asp Phe Arg
545 550 555 560
Leu Gly His Cys Pro Val Leu Val Ala Thr Ser Val Ala Ala Arg Gly
565 570 575
Leu Asp Ile Glu Gln Val Gln His Val Val Asn Phe Asp Met Pro Ser
580 585 590
Ser Ile Asp Glu Tyr Val His Arg Ile Gly Arg Thr Gly Arg Cys Gly
595 600 605
Asn Thr Gly Arg Ala Val Ser Phe Phe Asn Pro Glu Ser Asp Thr Pro
610 615 620
Leu Ala Arg Ser Leu Val Lys Val Leu Ser Gly Ala Gln Gln Val Val
625 630 635 640
Pro Lys Trp Leu Glu Glu Val Ala Phe Ser Ala His Gly Thr Thr Gly
645 650 655
Phe Asn Pro Arg Gly Lys Val Phe Ala Ser Thr Asp Ser Arg Lys Gly
660 665 670
Gly Ser Phe Lys Ser Asp Glu Pro Pro Pro Ser Gln Thr Ser Ala Pro
675 680 685
Ser Ala Ala Ala Ala Ala Asp Asp Glu Glu Trp Glu
690 695 700

<210> 7
<211> 661
<212> PRT
<213> Drosophila Melanogaster

<400> 7
Met Ser Asp Asp Trp Asp Asp Glu Pro Ile Val Asp Thr Arg Gly Ala
1 5 10 15
Arg Gly Gly Asp Trp Ser Asp Asp Glu Asp Thr Ala Lys Ser Phe Ser
20 25 30
Gly Glu Ala Glu Gly Asp Gly Val Gly Gly Ser Gly Gly Glu Gly Gly
35 40 45
Gly Tyr Gln Gly Gly Asn Arg Asp Val Phe Gly Arg Ile Gly Gly Gly
50 55 60
Arg Gly Gly Gly Ala Gly Gly Tyr Arg Gly Gly Asn Arg Asp Gly Gly
65 70 75 80
Gly Phe His Gly Gly Arg Arg Glu Gly Glu Arg Asp Phe Arg Gly Gly
85 90 95
Glu Gly Gly Phe Arg Gly Gly Gln Gly Gly Ser Arg Gly Gly Gln Gly
100 105 110
Gly Ser Arg Gly Gly Gln Gly Gly Phe Arg Gly Gly Glu Gly Gly Phe
115 120 125
Arg Gly Arg Leu Tyr Glu Asn Glu Asp Gly Asp Glu Arg Arg Gly Arg
130 135 140
Leu Asp Arg Glu Glu Arg Gly Gly Glu Arg Arg Gly Arg Leu Asp Arg
145 150 155 160
Glu Glu Arg Gly Gly Glu Arg Gly Glu Arg Gly Asp Gly Gly Phe Ala
165 170 175
Arg Arg Arg Arg Asn Glu Asp Asp Ile Asn Asn Asn Asn Asn Ile Ala
180 185 190
Glu Asp Val Glu Arg Lys Arg Glu Phe Tyr Ile Pro Pro Glu Pro Ser
195 200 205
Asn Asp Ala Ile Glu Ile Phe Ser Ser Gly Ile Ala Ser Gly Ile His

```
     210                 215                 220
Phe Ser Lys Tyr Asn Asn Ile Pro Val Lys Val Thr Gly Ser Asp Val
225                 230                 235                 240
Pro Gln Pro Ile Gln His Phe Thr Ser Ala Asp Leu Arg Asp Ile Ile
                245                 250                 255
Ile Asp Asn Val Asn Lys Ser Gly Phe Lys Ile Pro Thr Pro Ile Gln
            260                 265                 270
Lys Cys Ser Ile Pro Val Ile Ser Ser Gly Arg Asp Leu Met Ala Cys
        275                 280                 285
Ala Gln Thr Gly Ser Gly Lys Thr Ala Ala Phe Leu Leu Pro Ile Leu
    290                 295                 300
Ser Lys Leu Leu Glu Asp Pro His Glu Leu Glu Leu Gly Arg Pro Gln
305                 310                 315                 320
Val Val Ile Val Ser Pro Thr Arg Glu Leu Ala Ile Gln Ile Phe Asn
                325                 330                 335
Glu Ala Arg Lys Phe Ala Phe Glu Ser Tyr Leu Lys Ile Gly Ile Val
            340                 345                 350
Tyr Gly Gly Thr Ser Phe Arg His Gln Asn Glu Cys Ile Thr Arg Gly
        355                 360                 365
Cys His Val Val Ile Ala Thr Pro Gly Arg Leu Leu Asp Phe Val Asp
    370                 375                 380
Arg Thr Phe Ile Thr Phe Glu Asp Thr Arg Phe Val Val Leu Asp Glu
385                 390                 395                 400
Ala Asp Arg Met Leu Asp Met Gly Phe Ser Glu Asp Met Arg Arg Ile
                405                 410                 415
Met Thr His Val Thr Met Arg Pro Glu His Gln Thr Leu Met Phe Ser
            420                 425                 430
Ala Thr Phe Pro Glu Glu Ile Gln Arg Met Ala Gly Glu Phe Leu Lys
        435                 440                 445
Asn Tyr Val Ser Val Ala Ile Gly Ile Val Gly Gly Ala Cys Ser Asp
    450                 455                 460
Val Lys Gln Thr Ile Tyr Glu Val Asn Lys Tyr Ala Lys Arg Ser Lys
465                 470                 475                 480
Leu Ile Glu Ile Leu Ser Glu Gln Ala Asp Gly Thr Ile Val Phe Val
                485                 490                 495
Glu Thr Lys Arg Gly Ala Asp Phe Leu Ala Ser Phe Leu Ser Glu Lys
            500                 505                 510
Glu Phe Pro Thr Thr Ser Ile His Gly Asp Arg Leu Gln Ser Gln Arg
        515                 520                 525
Glu Gln Ala Leu Arg Asp Phe Lys Asn Gly Ser Met Lys Val Leu Ile
    530                 535                 540
Ala Thr Ser Val Ala Ser Arg Gly Leu Asp Ile Lys Asn Ile Lys His
545                 550                 555                 560
Val Ile Asn Tyr Asp Met Pro Ser Lys Ile Asp Asp Tyr Val His Arg
                565                 570                 575
Ile Gly Arg Thr Gly Cys Val Gly Asn Asn Gly Arg Ala Thr Ser Phe
            580                 585                 590
Phe Asp Pro Glu Lys Asp Arg Ala Ile Ala Ala Asp Leu Val Lys Ile
        595                 600                 605
Leu Glu Gly Ser Gly Gln Thr Val Pro Asp Phe Leu Arg Thr Cys Gly
    610                 615                 620
Ala Gly Gly Asp Gly Gly Tyr Ser Asn Gln Asn Phe Gly Gly Val Asp
625                 630                 635                 640
Val Arg Gly Arg Gly Asn Tyr Val Gly Asp Ala Thr Asn Val Glu Glu
                645                 650                 655
Glu Glu Gln Trp Asp
            660
```

<210> 8
<211> 28
<212> DNA
<213> Homo Sapiens

<400> 8
ctctgcatca aaaccacaga cttgaagg 28

<210> 9
<211> 30
<212> PRT
<213> Homo Sapiens

<400> 9
Cys Glu Asp Asn Pro Thr Arg Asn Arg Gly Phe Ser Lys Arg Gly Gly
1 5 10 15
Tyr Arg Asp Gly Asn Asn Ser Glu Ala Ser Gly Pro Tyr Arg
20 25 30

<210> 10
<211> 29
<212> PRT
<213> Homo Sapiens

<400> 10
Val Asp Thr Arg Lys Gly Lys Ser Thr Leu Asn Thr Ala Gly Phe Ser
1 5 10 15
Ser Ser Arg Ala Pro Asn Pro Val Asp Asp Glu Ser Trp
20 25

<210> 11
<211> 20
<212> DNA
<213> Homo Sapiens

<400> 11
aagtcaccat gggggatgaa 20

<210> 12
<211> 28
<212> DNA
<213> Homo Sapiens

<400> 12
ttaagatctt ttttgcacag gagtaagc 28

<210> 13
<211> 21
<212> DNA
<213> Homo Sapiens

<400> 13
tgcatcaaaa ccacagactt g 21

<210> 14
<211> 20
<212> DNA
<213> Homo Sapiens

<400> 14
aatgccatca aaggaacagc 20

<210> 15
<211> 2172
<212> DNA
<213> Homo Sapiens

<400> 15
atgggggatg aagattggga agcagaaatc aaccctcata tgtcttccta tgttcccata 60
tttgagaagg ataggtattc tggagaaaat ggagacaatt ttaacaggac tccagcttca 120
tcatcagaaa tggatgatgg accttctcga agagatcatt tcatgaaaag tggatttgcc 180
tctgggcgga attttggaaa cagagatgct ggtgagtgta ataagcgaga taatacatcc 240
acaatgggtg gttttggagt tggaaagagt tttggaaaca gaggtttttc aaacagcagg 300
tttgaagatg gtgatagctc tggtttctgg agagagtcta gtaatgactg cgaagataat 360
ccaacacgga acagagggtt ttccaagaga ggcggctatc gagatggaaa taattcagaa 420
gcttcagggc catacagaag aggtggaaga ggtagtttcc gaggttgccg tggaggattt 480
ggtctaggaa gtccaaataa tgacttagac ccagacgaat gtatgcagcg cactggtggc 540
ctttttggtt ctagaagacc agtattaagt ggcacaggta atggtgatac ttctcaaagc 600
agaagtggca gtggaagtga acgaggtggt tacaaaggtt taaatgaaga agtaataaca 660
ggctctggaa agaattcttg gaagtcagaa gcagaaggag gagaaagtag tgatactcaa 720
ggaccaaaag tgacctacat acccccctcct ccacctgagg atgaggactc catctttgca 780
cattatcaga caggcataaa cttcgacaaa tacgacacta ttcttgtgga agtgtctgga 840
catgatgcac caccagcaat tctgactttt gaagaagcta atctctgtca gacactgaat 900
aacaacattg ctaaagctgg ttatactaag cttactcctg tgcaaaaata cagtattcct 960
atcatacttg caggacgaga tttgatggct tgtgctcaaa cagggtctgg gaagactgcg 1020
gcttttctcc taccaatttt ggctcatatg atgcatgatg gaataactgc cagtcgtttt 1080
aaagagttgc aggaaccaga gtgtattatt gtagcaccaa ctcgagaatt ggtcaaccag 1140
atttatttgg aagccagaaa attttctttt gggacttgtg taagagctgt tgttatatat 1200
gggggaaccc agctgggaca ttcaattcga caaatagtac aaggctgtaa tatattatgt 1260
gctactcctg gaagactgat ggatatcata ggcaaagaaa agattggtct caaacagatc 1320
aaatacttag ttttggatga agctgatcgc atgttggata tgggttttgg tccagaaatg 1380
aagaagttaa tttcttgccc aggaatgcca tcaaaggaac agcgccaaac ccttatgttc 1440
agtgcaactt ttccagagga aattcaaagg ttggctgcag agtttttaaa gtcaaattat 1500
ctgtttgttg ctgttggaca agtgggtgga gcatgtagag atgttcagca gaccgttctc 1560
caagttggcc agttctcaaa aagagaaaag ctcgttgaaa ttctgcgaaa cataggggat 1620
gaaagaacta tggtctttgt tgaaactaag aaaaaagcag attttactgc aacttttctt 1680
tgtcaagaaa aaatatcaac tacaagtatc catggtgatc gggaacagag agagcgggag 1740
caagctcttg gagattttcg ctttggaaag tgcccagttc ttgttgctac ttcagtagct 1800
gccagagggc tggatattga aaatgtgcaa catgttatca attttgatct tccttctacc 1860
attgatgaat atgttcatcg aattgggcgt actggtcgtt gtgggaatac tggcagagca 1920
atttcctttt ttgatcttga atcggataac catttagcac agcctctagt aaaagtattg 1980
acagatgctc aacaggatgt tcctgcatgg ttggaagaaa ttgcctttag tacatacatt 2040
cctggcttca gtggtagtac aagaggaaac gtgtttgcat cagttgatac cagaaagggc 2100
aagagcactt tgaacacagc tgggttttct tcttcacgag ctcccaatcc agtagatgat 2160
gagtcatggg at 2172